ICASE

VERY HIGH ORDER ACCURATE TVD SCHEMES

Stanley Osher

and

Sukumar Chakravarthy

Report No. 84-44

September 10, 1984

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

Very High Order Accurate TVD Schemes

by

Stanley Osher*

Mathematics Department
University of California
Los Angeles, CA 90024

and

Sukumar Chakravarthy**

Rockwell Science Center
1049 Camino Dos Rios
Thousand Oaks, CA 91360

*Research supported in part by NSF Grant No. MCS 82-00788, ARO Grant No. DAAG 29-82-0090, NASA Grant No. NAG-1-270, NASA Consortium Agreement No. NCA2-1R390-403 and in part by the National Aeronautics and Space Administration under Contract No. NAS1-17070 while the author was in residence at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23665.

**Research supported by NASA Grant No. NAG-1-269.

Abstract

A systematic procedure for constructing semi-discrete families of 2m-1 order accurate, 2m order dissipative, variation diminishing, 2m+1 point band width, conservation form approximations to scalar conservation laws is presented. Here *m* is any integer between 2 and 8. Simple first order forward time discretization, used together with any of these approximations to the space derivatives, also results in a fully discrete, variation diminishing algorithm. These schemes all use simple flux limiters, without which each of these fully discrete algorithms is even linearly unstable. Extensions to systems, using a nonlinear field-by-field decomposition are presented, and shown to have many of the same properties as in the scalar case. For linear systems, these nonlinear approximations are variation diminishing, and hence convergent. A new and general criterion for approximations to be variation diminishing is also given. Finally, numerical experiments using some of these algorithms are presented.

Introduction

Recently there has been an enormous amount of activity related to the construction and analysis of "high resolution" schemes approximating hyperbolic systems of conservation laws. Some samples of the successful consequences of this activity can be found in the proceedings of the sixth AIAA Computational Fluid Dynamics Conference [3], [17], [31]. Extensive bibliographies can also be found in these papers.

Our aim here is to extend the use of these methods by making them even more accurate. We shall give a systematic procedure for constructing semi-discrete approximations to scalar conservation laws. Except for isolated critical parts, these schemes will have 2m-1 order accuracy, 2m order dissipation, and a bandwidth using 2m+1 points, for m any integer between two and eight. They are in conservation form and TVD - the variation of the discrete solution is non-increasing in time. Hence, no spurious oscillations are possible.

The high resolution schemes constructed earlier [13], [21], [28] use five points $(m = 2)$, and have second order accuracy. Some of these were proven to satisfy a single entropy inequality and hence to be convergent when $f(q)$ in (1.1) below is convex [20], [21]. It is possible that the piecewise parabolic method of Woodward and Colella [7], is third order accurate, and shares some of the properties discussed here when $m = 2$.

In a parallel work [4], we shall extend the construction below for $m = 2$, in order to approximate systems of conservation laws in multi-dimensions, using triangle-based algorithms. That work stresses the computational aspects of the algorithms, especially as they relate to the Euler equations of compressible gas dynamics.

Conventional schemes such as Lax-Wendroff even with an entropy fix [16] seem to lack a variation bound, although the convergence of this method for scalar convex $f(q)$ can now be proved (DiPerna, private communication). From a practical point of view, this lack of a variation bound seems to lead to a lack of robustness when computing complex flows with strong shock waves and steep gradients.

Another drawback of most finite-difference schemes is that discontinuities are approximated by discrete transitions, that when narrow, usually overshoot or undershoot, or when monotone, usually spread the discontinuity over many grid points.

Upwind schemes have been designed and used over the years, largely because of their success in treating this difficulty. Those based on solving the Riemann problem either exactly Godunov's method [9] or approximately, e.g. (Osher's [18], or Roe's [23] with an entropy fix [24], [4]), have been extremely successful, especially when put in a second-order accurate, high resolution framework, e.g. [3],[17], [31].

We should particularly mention the early investigations of van Leer [28], [29]. There he introduced the concepts of flux limiters, and higher order Riemann solvers. Recently Harten [13], using an argument also used in [1] and elsewhere, obtained sufficient conditions which he showed to be compatible with second order accuracy, and which guarantee that a scalar one-dimensional approximation is TVD - total variation diminishing. He constructed a scheme having that property and formally extended it to systems, using a field-by- field limiter, and Roe's decomposition.

We would also like to mention the work of Boris and Book [34], and Zalesak [32], concerning FCT schemes. They used flux limiters to supress oscillations in their schemes.

Harten's construction in [13] was done first for a fully discrete, explicit in time approximation. P. Sweby [26] has investigated the properties of various limiters in this context. We shall not use Sweby's ideas here since we seek higher order accuracy, and his symmetry restriction would make our approximations only second order accurate in the semi-discrete context.

We shall use the now-introduced term "high resolution scheme" to mean a formal extension to systems via a field-by-field decomposition, of a scalar, higher than first order accurate, variation diminishing scheme. These schemes do not, in general, satisfy the entropy condition - e.g. expansion shocks exist as stable solutions of high resolution schemes based on Roe's (unmodified) scheme. In [21] we used Osher's decomposition and certain limiters to prove that limit solutions of a class of second order accurate high resolution schemes do satisfy the entropy condition for hyperbolic systems of conservation laws. We also proved convergence of another class of high resolution approximations to scalar convex conservation laws in [20] as well as in [21]. We believe that the ideas concerning the entropy condition in these two papers can be extended to the high order accurate schemes constructed in the following sections, but we do not attempt this here. The interested reader might also consider the remarks on entropy fixes in [4], [19] and [16].

The high-order accurate TVD schemes are first obtained here for semi-discrete (continuous in time) approximations, and can thus serve as a guideline for a wide variety of time discretizations, both implicit and explicit. See [2] for efficient implicit calculations approximating Euler's equations in transonic and supersonic aeronautics. TVD schemes also have a certain diagonal dominance that is very useful in implicit methods [2], [12],

An interesting and useful fact concerning time discretization is the following (mentioned in Theorems (3.1) and (3.2) below). All of the semi-discrete approximations constructed below are unconditionally (even linearly) unstable when (a) they are used together with simple first order accurate forward Euler time discretizations, and (b) the flux limiters are removed. However, they are all conditionally stable when the limiters, which enforce the variation bound, are kept. Thus, although the limiters might not act at all on a resulting steady state solution, they act non-linearly during transient calculations to enforce the variation bound. This elementary time differencing is sometimes useful, e.g. when steady state calculations on coarse grids are to be obtained simply.

Goodman and Leveque have recently shown [10] that two space dimensional scalar approximations cannot be TVD and still be more than first order accurate, given that the associated flux functions are reasonably smooth. Nevertheless two dimensional schemes based on dimension by dimension TVD differencing have worked quite well, even for complex configurations with very strong shocks. See e.g. [3], [5], [7]. In particular, it seems that our remark in the previous paragraph about conditional stability of Euler forward time discretization is also experimentally valid here. Perhaps a more sophisticated, scheme dependent, notion of variation is needed for the theory in several space dimensions.

The format of this paper is as follows. In section I, we review the relevant theory of weak solutions of conservation laws and their approximations. In section II, we exemplify our general theory by constructing families of second and third order accurate TVD schemes using five points. In section III, we perform the general construction for scalar conservation laws and state Theorems (3.1) and (3.2) which contain the main results of this paper. In section IV we prove the theorems. In section V we obtain an apparently new and general criterion for an approximation to be TVD, which we hope will be useful. In section VI we extend our construction to high resolution schemes approximating systems. Section VII contains some numerical evidence demonstrating the utility of these methods. Many more experimental results are given in [5].

I. Review of Theory of Weak Solutions and their Approximations

We shall consider numerical approximations to the initial value problem for nonlinear hyperbolic systems of conservation laws.

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} f(q) = 0,\ t > 0,\ -1 \leq x < 1 \tag{1.1}$$

with periodic boundary conditions:

$$q(x+1,t) = q(x,t),$$

given initial conditions $q(x,0)$.

Here $q(x,t)$ is an m-vector of unknowns, and the flux function $f(q)$ is vector-valued, having m components. The system is hyperbolic when the Jacobian matrix has real eigenvalues.

It is well-known that solutions of (1.1) may develop discontinuities in finite time, even when the initial data are smooth. Because of this, we seek a weak solution of (1.1).

These weak solutions are not necessarily unique. For physical reasons, the limit of the viscous equation, as viscosity tends to zero is sought. This leads to an infinite family of inequalities in the scalar case which when satisfied by so-called "entropy" solutions to (1.1) yield well-posedness in L^1 of the evolution problem. This result is due to Kruzkov [15].

For systems of equations, Lax has defined an entropy inequality using an entropy function [35]. The entropy inequality satisfied by "entropy" solutions to systems has an important geometric consequence concerning admissible discontinuities.

This theory is quite well developed and often reviewed - see e.g. [19], section II. One new result is the following; in the scalar convex case, a single entropy inequality is equivalent to the required infinite number, if the solution is of bounded variation. (See [8].) This fact was crucial to the convergence results in [20] and [21].

Next we consider a semi-discrete, method of lines, approximation to (1.1). We break the interval $(-1,1)$ into subintervals:

$$I_j = \{x/(j-\frac{1}{2})\Delta x \le x \le (j+\frac{1}{2})\Delta x\}$$

$j = 0,\pm 1,\ldots,\pm N$, with $(2N+1)\Delta x = 2$

Let $x_j = j\Delta x$, be the center of each interval I_j, with end points $x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}$.

Define the step function for each $t > 0$, as

$$Q_{\Delta x}(x,t) = q_j(t),$$

for $x \in I_j$.

The initial data is discretized via the averaging operator $T_{\Delta x}$,

$$T_{\Delta x}q(x,0) = \frac{1}{\Delta x}\int_{I_j} q_o(s,0)ds = q_j(o), \text{ for } x \in I_j$$

For any step function, we define the difference operators

$$\Delta_{\pm} g_j = \pm(q_{j\pm 1} - q_j)$$

$$D_{\pm} g_j = \frac{1}{\Delta x}\Delta_{\pm} g_j$$

A method of lines, conservation form, discretization of (1.1), is a system of differential equations

$$\frac{\partial}{\partial t}q_j + D_{+}\hat{f}_{j-\frac{1}{2}} = 0,\ j = 0,\pm 1,\ldots,\pm N. \tag{1.2}$$

$$Q_{\Delta}(x,0) = T_{\Delta x}q(x,0) \text{ for } x \in I_j$$

Here, the numerical flux defined by:

$$\hat{f}_{j-\frac{1}{2}} = \hat{f}(q_{j+k-1}, \ldots, q_{j-k}), \tag{1.3}$$

for $k \geq 1$, is a Lipschitz continuous function of its arguments, satisfying the consistency condition:

$$\hat{f}(q,q,\ldots,q) = f(q)$$

It is well known that bounded a.e. limits as $\Delta x \to 0$, of approximate solutions converge to weak solutions of (1.1). This does not necessarily imply that limit solutions will satisfy any of the above- mentioned entropy conditions. Some restrictions on the numerical flux are required.

The most general class of scalar flux functions known to yield convergent approximations whose limit solutions will satisfy all entropy conditions, for general scalar $f(q)$ is the class of "E" fluxes, introduced in [18].

A consistent scheme whose numerical flux $h_{j-\frac{1}{2}}$ satisfies

$$sgn\ (q_j - q_{j-1})[h_{j-\frac{1}{2}} - f(q)] \leq 0 \tag{1.4}$$

for all q between q_{j-1} and q_j, is said to be an E flux.

Other equivalent definitions are given in [18] and [27].

Unfortunately these schemes are at most first order accurate [18]. We shall use three point E schemes as building blocks for our higher order accurate TVD schemes described in the next sections. We have already done this to get convergent, second order TVD schemes, approximating convex scalar conservation laws in [20], and [21].

Examples of three-point E schemes include three-point monotone schemes, e.g. Engquist-

Osher's [33], Godunov's [9] (which is canonical - see [18]), or entropy fixes of Roe's scheme [24]. These are defined again at the end of this section.

Together with an entropy inequality, a key estimate involved in most convergence proofs is a bound on the variation. For any fixed $t \geq 0$, the x variation of scalar $Q_{\Delta x}(x,t)$ is defined as

$$B(Q_{\Delta x}) = \sum_j |\Delta_+ q_j|$$

If we can write for every

$$\Delta_+ \hat{f}_{j-\frac{1}{2}} = -C_{j+\frac{1}{2}}\, \Delta_{+,}q_j + D_{j-\frac{1}{2}} \Delta_- q_j \tag{1.5a}$$

$$C_{j+\frac{1}{2}} \geq 0 \tag{1.5b}$$

$$D_{j-\frac{1}{2}} \geq 0 \tag{1.5c}$$

then it is easy to show, [21], using an argument of [25], that for $t_1 \geq t_2 \geq 0$.

$$B(Q_{\Delta x}(\cdot,t_1)) \leq B(Q_{\Delta x}(\cdot,t_2)) \tag{1.6}$$

Harten in [13], pointed out for explicit methods, that this decomposition could be obtained for schemes which are higher than first order accurate. See also earlier work by van Leer [28]. In section V we obtain a more general criterion than (1.5), guaranteeing that (1.6) is valid. We shall use criterion (1.5) here to get very high order accurate, TVD schemes of the type

$$\frac{\partial}{\partial t} q_j = -\Delta_+ \hat{f}_{j-\frac{1}{2}} = C_{j+\frac{1}{2}} \Delta_+ q_j - D_{j-\frac{1}{2}} \Delta_- q_j \tag{1.7a}$$

with

$$C_{j+\frac{1}{2}} = C(q_{j+m}, \ldots, q_{j-m+1}) \geq 0 \tag{1.7b}$$

$$D_{j-\frac{1}{2}} = D(q_{j+m-1}, \ldots, q_{j-m}) \geq 0 \tag{1.7c}$$

which are $2m-1$ order accurate, except at isolated critical points, for $2 \leq m \leq 8$.

In addition to (1.6) we have a maximum principle for (1.7)

$$\min_k q_k(0) \le q_j(t) \le \max_k q_k(0), \tag{1.8}$$

for each j and all $t \ge 0$, [21].

Moreover, in [21], we also showed a limit on the possible accuracy of approximations of type (1.17), for $m = 2$. A glance at the proof of that Lemma (2.3) shows that the result is also valid for general m, namely:

Approximation (1.7) is at most first order accurate at nonsonic critical points of q, i.e. points $\bar{q}$ at which $f'(\bar{q}) \neq 0 = \bar{q}_x$.

In spite of this local degeneracy, higher order accuracy, combined with TVD does improve performance, even when discontinuities are present. This is shown numerically in ref [5] and elsewhere.

As promised, we now present several useful three-point E fluxes.

Engquist-Osher

$$\begin{aligned} h^{EO}(q_j,q_{j-1}) = &\int_0^{q_j} \min(f'(s),0)ds \\ &+\int_0^{q_{j-1}} \max(f'(s),0)ds + f(0) \end{aligned} \tag{1.9}$$

Godunov

$$h^G(q_j,q_{j-1}) = \min_{q_{j-1} \le q \le q_j} f(q), \text{ if } q_{j-1} \le q_j \tag{1.10}$$

$$= \max_{q_{j-1} \ge q \ge q_j} f(q), \text{ if } q_{j-1} > q_j$$

Roe with entropy fix, approximating a convex $f(q)$ i.e., $f'' \ge 0$ with $f'(q) = 0$ at a single sonic point $\bar{q}$. Define

$$h^{RV}(q_j,q_{j-1}) = \frac{1}{2}[(f(q_j) + f(q_{j-1})) - \left|\frac{\Delta_- f(q_j)}{\Delta_- q_j}\right|\Delta_- q_j] \qquad (1\text{-}11)$$

unless

$$q_{j-1} < \bar{q} < q_j,$$

then take any Lipschitz function so that:

$$h^{RV}(q_j,q_{j-1}) \le f(\bar{q})$$

See e.g. [4], [24], for various fixes of this type.

II. Second and Third Order Accurate TVD Schemes
Which Use a Five-Point Module

We begin by exemplifying our general theory using a very important and convenient class of schemes. We shall approximate the scalar conservation law by a family of five point, semi-discrete method of lines, and TVD approximations.

Let $h(q_{j+1},q_j)$ be the numerical flux corresponding to a three- point E scheme. Next we define

$$df^-_{j+\frac{1}{2}} = h(q_{j+1},q_j) - f(q_j) \tag{2.1a}$$

$$df^+_{j+\frac{1}{2}} = f(q_{j+1}) - h(q_{j+1},q_j) \tag{2.1b}$$

We can then write

$$h(q_{j+1},q_j) = \frac{1}{2}\Big[f(q_{j+1}) + f(q_j)\Big] - \frac{1}{2}\Big[df^+_{j+\frac{1}{2}} - df^-_{j+\frac{1}{2}}\Big]$$

These new quantities df^- and df^+ denote the difference in flux across the waves with negative and positive velocities respectively in the interval under consideration. The subscript $_{j+\frac{1}{2}}$ denotes the interface between two cells whose centroids are denoted by grid points with subscripts j and $j+1$ respectively. Thus $df^+_{j+\frac{1}{2}}$ denotes the difference (taken from right to left) in flux across all the positive (forward) breaking waves at the cell interface $_{j+\frac{1}{2}}$, etc.

A general semi-discrete conservation form approximation to (1.1) can be given as

$$q_t + \frac{\left(\hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}}\right)}{\Delta x} = 0 \tag{2.2}$$

Here the quantity $\hat{f}$ is the representative for numerical flux.

With this notation the numerical flux of one new family of TVD schemes can be represented by

$$\hat{f}_{j+\frac{1}{2}} = h(q_{j+1},q_j) - \alpha(df^-_{j+3/2})^{(1)} - (\frac{1}{2}-\alpha)(df^-_{j+\frac{1}{2}})^{(0)} \tag{2.3}$$
$$+ (\frac{1}{2}-\alpha)(df^+_{j+\frac{1}{2}})^{(0)} + \alpha(df^+_{j-\frac{1}{2}})^{(-1)},$$

for $0 < \alpha \leqslant \frac{1}{2}$.

The superscripts shown over the df denote flux-limited values of df, and are computed as follows:

$$\left(df^-_{j+3/2}\right)^{(1)} = \text{min mod}\left[df^-_{j+3/2},\ b\ df^-_{j+\frac{1}{2}}\right] \tag{2.4a}$$

$$\left(df^-_{j+\frac{1}{2}}\right)^{(0)} = \text{min mod}\left[df^-_{j+\frac{1}{2}},\ b\ df^-_{j+3/2}\right] \tag{2.4b}$$

$$\left(df^+_{j+\frac{1}{2}}\right)^{(0)} = \text{min mod}\left[df^+_{j+\frac{1}{2}}, b\ df^+_{j-\frac{1}{2}}\right] \tag{2.4c}$$

$$\left(df^+_{j-\frac{1}{2}}\right)^{(-1)} = \text{min mod}\left[df^+_{j-\frac{1}{2}}, b\ df^+_{j+\frac{1}{2}}\right] \tag{2.4d}$$

In the above, the operator "min mod" is defined by:

$$\text{min mod}\,[x,y] = (sgn\ x)\ \text{max}(0,\text{min}|x|,y\ \text{sgn}\ x) \tag{2.5}$$

(see e.g. [26]), and b is a "compression" parameter chosen in the range

$$1 < b \le 1 + \frac{1}{2\alpha} = b_{\max}. \tag{2.6}$$

The case $\alpha = 0$ also yields a TVD scheme, but this one is not time dissipative, so steady state solutions are difficult to obtain. We recommend that α be positive in all applications. The dissipation in our general algorithm is an increasing function of α.

The non-TVD or unlimited forms of the schemes in the new family can be obtained by replacing the $(df)^{(v)}$ terms appearing in (2.3) with the corresponding unlimited df values. The truncation error of the unlimited form (up to second order) is given by:

$$TE = (\frac{1}{6} - \alpha)(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u) \tag{2.7}$$

It is interesting to note that TE is independent of the particular E-scheme used, i.e. independent of h.

Particular schemes in the new family may be chosen by picking various values for the parameter α. Some special cases are summarized in Table 2.1. The TE shown in the last column corresponds to the unlimited forms. The names given to the TVD schemes are based on the names used in the literature, e.g. [29], for the corresponding unlimited schemes.

Value of α	Name of TVD Scheme	$b_{\max}$	2nd order TE
1/6	Third-Order	4	0
1/2	Fully Upwind	2	$-1/3(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$
1/4	Fromm's	3	$\frac{-1}{12}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$
1/8	Low TE second-order	5	$\frac{1}{24}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$
0	Central	∞	$\frac{1}{6}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$
1/3	No Name	5/2	$\frac{-1}{6}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$

Table 2.1 Particular Cases of New Family of TVD Schemes

Semi-discrete notions of TVD schemes only show that, when a suitable time discretization is chosen, the overall algorithm is TVD, hence has a convergent subsequence as $\Delta t \to 0$. See e.g. [21]. There is always a CFL restriction on explicit schemes. For simplicity, we consider the explicit scheme given by forward Euler time discretization:

$$\frac{q_j^{n+1}-q_j^n}{\Delta t} + \frac{\hat{f}^n_{j+\frac{1}{2}} - \hat{f}^n_{j-\frac{1}{2}}}{\Delta x} = 0 \tag{2.8}$$

This is only first order accurate in time.

As part of a general result Theorem (3.2), it follows that the *unlimited* versions of (2.8) are all *unstable* for any CFL number $\lambda = \frac{\Delta t}{\Delta x}$. However, Table (2.2) gives *stable* time steps for the *flux-limited* versions for $b = b_{\max}$.

The general condition for (2-8) to be TVD is:

$$\frac{\Delta t}{\Delta x}\left(\frac{df^+_{j+\frac{1}{2}} - df^-_{j+\frac{1}{2}}}{\Delta_+ q_j}\right) \le \frac{4\alpha}{1+4\alpha} \quad \text{if } b = b_{\max}. \tag{2.9a}$$

and for general b it is

$$\frac{\Delta t}{\Delta x}\left(\frac{df^+_{j+\frac{1}{2}} - df^-_{j+\frac{1}{2}}}{\Delta_+ q_j}\right) \le \frac{1}{1+\alpha+b(\frac{1}{2}-\alpha)} \tag{2.9b}$$

Value of α	b_{max}	$\left(\frac{\Delta t}{\Delta x}\left(\frac{df^{+}_{j+\frac{1}{2}} - df^{-}_{j+\frac{1}{2}}}{\Delta_{+}q_j}\right)\right)_{max}$
1/6	4	2/5
1/2	2	2/3
1/4	3	1/2
1/8	5	1/3
0	∞	0
1/3	5/2	4/7

TABLE 2.2. Stable Time Steps for This New Class of TVD Schemes.

In equations (2.4a) and (2.4b) the flux-limited values of df are defined. This value is computed in some interval by comparing the original unlimited value with its neighboring value, after that neighbor has been multiplied by the "compression" parameter b. Assuming that the two values being compared are of the same sign, the "min mod" operator chooses the one whose absolute value is the smallest. If $b > 1$, the flux-limited value returned most often will be the unlimited value itself. Thus, for most grid points (away from high second-gradient regions where the unlimited value of slope df can be much greater than the unlimited value of the neighboring slope), the TVD scheme is identical to the corresponding unlimited scheme. (Having a larger value of b enhances this property.) At critical points of the fluxes, the neighboring values of df can be of opposite sign. There, the "min mod" operator returns the value zero. Thus, away from maxima, minima, and points of discontinuity, the TVD scheme reduces to its corresponding unlimited scheme.

We next present a class of schemes having the same five-point band width and which are all third-order accurate in their unlimited versions. However, the flux limiting is a bit more far-reaching than in the α class defined above. This may cause a slight deterioration of accuracy when we use a coarse grid to approximate solutions having many critical points.

The flux is defined by:

$$\begin{aligned}\hat{f}_{j+\frac{1}{2}} = h(q_{j+1},q_j) &- (\tfrac{1}{12}+\beta)(df^{-}_{j+\frac{3}{2}})^{(1)} \\ &- (\tfrac{1}{2}-2\beta)(df^{-}_{j+\frac{1}{2}})^{(0)} \\ &+ (\tfrac{1}{12}-\beta)(df^{-}_{j-\frac{1}{2}})^{(-1)} \\ &- (\tfrac{1}{12}-\beta)(df^{+}_{j+\frac{3}{2}})^{(1)} \\ &+ (\tfrac{1}{2}-2\beta)(df^{+}_{j+\frac{1}{2}})^{(0)} \\ &+ (\tfrac{1}{12}+\beta)(df^{+}_{j-\frac{1}{2}})^{(-1)}\end{aligned} \tag{2.10}$$

Here we take $0 < \beta \leq \frac{1}{12}$. Again, $\beta = 0$ corresponds to a non-dissipative, central difference, but TVD scheme.

The flux-limited values of df are defined through:

$$(df^{-}_{j+\frac{3}{2}})^{(1)} = \min\bmod\left[df^{-}_{j+\frac{3}{2}}, b\,df^{-}_{j+\frac{1}{2}}\right] \tag{2.11a}$$

$$(df^{-}_{j+\frac{1}{2}})^{(0)} = \min\bmod\left[df^{-}_{j+\frac{1}{2}}, b\,df^{-}_{j+\frac{3}{2}}\right] \tag{2.11b}$$

$$(df^{-}_{j-\frac{1}{2}})^{(-1)} = \min\bmod\left[df^{-}_{j-\frac{1}{2}}, b\,df^{-}_{j+\frac{1}{2}}, b\,df^{-}_{j+3/2}\right] \tag{2.11c}$$

$$(df^+_{j+3/2})^{(1)} = \min \bmod \left[df^+_{j+\frac{3}{2}}, bdf^+_{j+\frac{1}{2}}, bdf^+_{j-\frac{1}{2}}\right] \tag{2.11d}$$

$$(df^+_{j+\frac{1}{2}})^{(0)} = \min \bmod \left[df^+_{j+\frac{1}{2}}, bdf^+_{j-\frac{1}{2}}\right] \tag{2.11e}$$

$$(df^+_{j-\frac{1}{2}})^{(-1)} = \min \bmod \left[df^+_{j-\frac{1}{2}}, bdf^+_{j+\frac{1}{2}}\right] \tag{2.11f}$$

In the above, the operator "min mod" of three quantities is defined through

$$\min \bmod [x,y,z] = \min \bmod [\min \bmod [x,y],z], \tag{2.12}$$

This is easily seen to be independent of the order of x,y,z. Again, b is a "compression" parameter. Here it is chosen in the range.

$$1 < b \leq 3 + 12\beta \tag{2.13}$$

The non-TVD or unlimited forms of these schemes are obtained by replacing each $(df)^{(\nu)}$ term by its corresponding unlimited df value. The third order truncation error of the unlimited form coincides with the dissipation and is proportional to β. See the proof of Theorem (3.1) below.

For $\beta = 1/12$, this scheme (2.10),(2.11) coincides with (2.3),(2.4) for $\alpha = 1/6$. The limiting simplifies a bit here, since the coefficients of (2.11c) and (2.11d) vanish. For this reason, we prefer this scheme to any of the other third order "β" schemes.

Finally, we compute the CFL number guaranteeing that (2.8),(2.10),(2.11) is TVD. The results are:

(2.14a) (for general β satisfying (2.13)

$$\frac{\Delta t}{\Delta x} \frac{df^+_{j+\frac{1}{2}} - df^-_{j+\frac{1}{2}}}{\Delta_+ q_j} \leq \left[\frac{13}{12} + \beta + b(\frac{7}{12} - 3\beta)\right]^{-1}$$

(2.14b) (for $b = 3 + 12\beta = b_{max}$)

$$\frac{\Delta t}{\Delta x}\,\frac{df^{+}_{j+\frac{1}{2}} - df^{-}_{j+\frac{1}{2}}}{\Delta_{+}q_{j}} \leq \left[\frac{34}{12} - \beta - 36\beta^{2}\right]^{-1}$$

III. General 2m-1 and 2m-2 Order Accurate TVD Schemes

Which Use a 2m+1 Point Module for $m \le 8$.

We use the notation of the previous section to approximate (1.1) via a family of schemes of the type (2.2) where:

$$\hat{f}_{j+\frac{1}{2}} = \hat{f}^{m,\beta}_{j+\frac{1}{2}} = h(q_{j+1},q_j) + \sum_{k=-m+1}^{m-1}\left(\mu_k^m + (-1)^k\beta\binom{2m-2}{k+m-1}\right)\left(df^-_{j+k+\frac{1}{2}}\right)^{(k)} \tag{3.1}$$

$$+ \sum_{k=-m+1}^{m-1}\left(\nu_k^m - (-1)^k\beta\binom{2m-2}{k+m-1}\right)\left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}$$

Here m is an integer, $m \ge 2$, and β satisfies $0 < \beta < \left(m\binom{2m}{m}\right)^{-1}$. (The upper bound on β could be relaxed considerably, at a cost of complicating our calculations. We shall not do this here.) The binomial coefficient is defined for A,B integers with $0 \le B \le A$, as usual:

$$\binom{A}{B} = \frac{A!}{B!(A-B)!}$$

The coefficients ν_k^m, μ_k^m can be defined recursively by:

$$\nu^m_{m-1} = (-1)^{m-1}\left(m\binom{2m}{m}\right)^{-1} \quad ,for\ m \ge 2. \tag{3.2a}$$

$$\nu_0^m = \frac{1}{2} \tag{3.2b}$$

$$\nu_k^m = -\nu^m_{-k}, \quad k = 1,\ldots,m-1 \tag{3.2c}$$

and

$$\nu_k^{m+1} = \nu^m_{-k} + (-1)^k\left((m+1)\binom{2m+2}{m+1}\right)^{-1}\frac{k}{m}\binom{2m}{m-k} \quad for\ k = 1,2,\ldots,m-1. \tag{3.2d}$$

An alternative direct formulation comes by defining:

$$\nu_k^m = \sum_{j=k}^{m-1} \lambda_j^m = -\nu_{-k}^m \quad \mathit{for} \quad k = 1,\ldots,m-1, \tag{3.3a}$$

where

$$\lambda_k^m = \sum_{j=k+1}^{m} (-1)^{j-1} \left(j\binom{2m}{m}\right)^{-1} \binom{2m}{m+j}, \tag{3.3b}$$

and

$$\nu_0^m = \frac{1}{2} \tag{3.3c}$$

Also:

$$\mu_k^m = \nu_k^m, \quad \mathit{for} \quad k = \pm 1, \pm 2, \ldots, \pm(m-1), \tag{3.4a}$$

$$\mu_0^m = -\frac{1}{2} \tag{3.4b}$$

We define the flux limited quantities as follows. For each j:

$$\left[df^{-}_{j+\frac{1}{2}}\right]^{(k)} = \text{min mod}\left[df^{-}_{j+\frac{1}{2}}, bdf^{-}_{j-k+\frac{1}{2}}, bdf^{+}_{j-k+3/2}\right] \quad \text{for all k with } 0 \neq k \neq 1. \tag{3.5a}$$

$$\left[df^{-}_{j+\frac{1}{2}}\right]^{(0)} = \text{min mod}\left[df^{-}_{j+\frac{1}{2}}, bdf^{-}_{j+3/2}\right] \tag{3.5b}$$

$$\left[df^{-}_{j+\frac{1}{2}}\right]^{(1)} = \text{min mod}\left[df^{-}_{j+\frac{1}{2}}, bdf^{-}_{j-\frac{1}{2}}\right] \tag{3.5c}$$

$$\left[df^{+}_{j+\frac{1}{2}}\right]^{(k)} = \text{min mod}\left[df^{+}_{j+\frac{1}{2}}, bdf^{+}_{j-k+\frac{1}{2}}, bdf^{+}_{j-k-\frac{1}{2}}\right] \quad \text{for all k with } 0 \neq k \neq -1. \tag{3.5d}$$

$$\left[df^{-}_{j+\frac{1}{2}}\right]^{(0)} = \mathit{min\ mod}\left[df^{+}_{j+\frac{1}{2}}, bdf^{+}_{j-\frac{1}{2}}\right] \tag{3.5e}$$

$$\left[df^{+}_{j+\frac{1}{2}}\right]^{(-1)} = min\ mod \left[df^{+}_{j+\frac{1}{2}}, bdf^{+}_{j+\frac{3}{2}}\right] \tag{3.5f}$$

The compression parameter b is allowed to vary between:

$$0 < b < \left(\sum_{j=2}^{m} \frac{1}{2j-1}\right)^{-1} \left(1 + 2\beta \binom{2m-2}{m-1}\right) \tag{3.6}$$

We can now state the following:

Theorem (3.1) (*beta schemes*)

The scheme (2.2), (3.1)-(3.6) has the following properties:

(a) It is TVD and satisfies the maximum principle.

(b) For any $\beta, 0 < \beta \le \left(m\binom{2m}{m}\right)^{-1}$, and $m \le 7$, b can be taken to be greater than one. For $m = 8$, there exists β_o such that for $0 < \beta_o \le \beta \le \left(8\binom{16}{8}\right)^{-1}$, b can be again taken to be greater than one.

(c) The unlimited version is $(2m-1)-order$ accurate and $2m-order$ dissipative, with truncation error and dissipation both proportional to β. Thus, for $b > 1$, the TVD scheme will return $(2m-1)-order$ accuracy except at critical points, or points of discontinuity, where it is formally only first-order accurate.

(d) The simple Euler forward difference time discretized version (2.8) is TVD if the CFL restriction:

$$\frac{\Delta t}{\Delta x}\left(\frac{df^{+}_{j+\frac{1}{2}} - df^{-}_{j+\frac{1}{2}}}{\Delta_{+}q_j}\right) \le \Bigg[1+b\Bigg[\frac{1}{2}\sum_{j=2}^{m}\left(\frac{1}{2j-1}\right) + \frac{1}{2} - \sum_{j=2}^{m}\frac{1}{2j(2j-1)}$$

$$- \beta \binom{2m-2}{m-1} - \beta \binom{2m-2}{m-2} \Big]$$

$$+ \sum_{j=2}^{m} \frac{1}{2j(2j-1)} + \beta \binom{2m-2}{m-2} \Big]^{-1}$$

is valid.

(e) This same forward difference time-discretized scheme, without flux limiters, is linearly unstable for any CFL number.

This theorem will be proven in the next section.

These *beta* schemes give one more order of accuracy per $2m+1$ module than the *alpha -schemes* defined next, except for a special case $\alpha = 2\beta = 2\left(m\binom{2m}{m}\right)^{-1}$ when they coincide.

Again using the notation of the previous section, we approximate (1.1) via a family of schemes of the type (2.2), where

$$\hat{f}^{m,\alpha}_{j+\frac{1}{2}} = h(q_{j+1},q_j) + \sum_{k=-m+2}^{m-1} [\mu_k^{m-1} + (-1)^k \alpha \binom{2m-3}{k+m-2} \Big] \left(df^-_{j+k+\frac{1}{2}}\right)^{(k)} \tag{3.7}$$

$$+ \sum_{k=-m+1}^{m-2} \left[\nu_k^{m-1} - (-1)^k \alpha \binom{2m-3}{k+m-1}\right] \left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}$$

Here m is an integer with $m \geq 2$, and $0 < \alpha < \left((m-1)\binom{2m-2}{m-1}\right)^{(-1)}$. (Again, we impose the upper bound on α for simplicity only.) The coefficients ν_k^{m-1}, μ_k^{m-1} were defined in (3.2),(3.3),(3.4), and we also define

$$\nu^{m-1}_{-m+1} = 0 = \mu^{m-1}_{m-1} \tag{3.8}$$

The flux limited quantities are defined precisely as in (3.5), and now the quantity b is allowed to vary between

$$0 < b \leq \frac{1 + 2\alpha\binom{2m-3}{m-1}}{2\alpha\binom{2m-3}{m-1} + \sum_{j=2}^{m-1} (2j-1)^{-1}} \tag{3.9}$$

(where the sum $\sum_{j=2}^{1} (2j-1)^{-1} = 0$, by definition.)

We can now state the following:

Theorem 3.2 ("α" *schemes*).

The scheme defined via (2.2), (3.7) has the following properties

(a) It is TVD

(b) If $m \leq 8$, then b can be taken to be greater than one.

(c) If $\alpha = 2/m\binom{2m}{m}^{-1}$, then this scheme is identical to the "β" *scheme* for the same m, with $\beta = 1/m\binom{2m}{m}^{-1}$. Hence its unlimited version is $2m-1$-*order* accurate.

(d) For all other admissible values of α, the unlimited version of the scheme is $2m-2$ *order* accurate with $2m$ *order* dissipation, which is proportional to α. The truncation error is equal to

$$TE = (-1)^m\left[\alpha - 2/m\binom{2m}{m}^{-1}\right](\Delta x)^{2m-2}\left(\frac{\partial}{\partial x}\right)^{2m-1} f(u),$$

and is thus independent of the choice of h, the E flux.

(e) The simple Euler forward difference time discretized version, (2.8), is TVD if the CFL restriction:

$$
\begin{aligned}
\frac{\Delta t}{\Delta x}\left(\frac{\overrightarrow{df}_{j+\frac{1}{2}} - \overleftarrow{df}_{j+\frac{1}{2}}}{\Delta_{+}q_j}\right) \le \Bigg[1 + b\Bigg(\frac{1}{2} - \sum_{j=2}^{m-1}\frac{1}{2j(2j-1)} \\
+ \frac{1}{2}\sum_{j=1}^{m-2}\frac{1}{2j-1} - \alpha\binom{2m-3}{m-2}\Bigg] \\
+ \alpha\binom{2m-3}{m-2} + \sum_{j=2}^{m-1}\frac{1}{2j(2j-1)}\Bigg]^{-1}
\end{aligned}
$$

(f) This same forward difference time discretized scheme, without flux limiters, is linearly unstable for any CFL number.

This theorem will be proven in the next section.

IV. Proof of Main Theorems

Let the k^{th} power of the shift operator be defined as

$$S^k q_j = q_{j+k}$$

Define the central difference operator for $k=1,2,\ldots$

$$D_o(k\Delta x) = \frac{1}{2k\Delta x}\left(S^k - S^{-k}\right) \tag{4.1}$$

We shall use the following well-known formula - see e.g. [14]. Let q be any smooth function with $q_j = q(j\Delta x) = q(x)$. Then

$$q_x(x) = -2\sum_{k=1}^{m} \frac{(-1)^k (m!)^2}{(m+k)!(m-k)!} D_o(k\Delta x) q(x) \tag{4.2}$$

$$+ (-1)^m \frac{2(m!)(m+1)!}{(2m+2)!} (\Delta x)^{2m} \left(\frac{\partial}{\partial x}\right)^{2m+1} q(x) + 0((\Delta x))^{2m+1}$$

Let C^{2m} denote the operator from which uniquely defines $2m^{th}$ *order* accurate differencing based on central difference operators using a module $(-m,m)$. We define

$$C^{2m} = -2\sum_{k=1}^{m} (-1)^k \binom{2m}{m-k} \binom{2m}{m}^{-1} D_o(k\Delta x) \tag{4.3a}$$

$$= \sum_{k=-m}^{m-1} \lambda_k^m D_+ S^k$$

where

$$\lambda_k^m = \sum_{j=max(-k,k+1)}^{m} (-1)^{j+1} \binom{2m}{m-k} \left(j\binom{2m}{m}\right)^{-1} \tag{4.3b}$$

For our purposes, a better formulation is

$$C^{2m} = D_+S^{-1} + \sum_{k=-m+1}^{m-1} \nu_k^m D_+(\Delta_+S^{k-1}) \tag{4.4a}$$

with

$$\nu_k^m = \sum_{j=k}^{m-1} \lambda_j^m, \quad k=1,\ldots,m-1 \tag{4.4b}$$

$$\nu_o^m = \frac{1}{2} \tag{4.4c}$$

$$\nu_k^m = -\nu_{-k}^m, k = -1,\ldots,-m+1 \tag{4.4d}$$

It is fairly simple to verify (4.3) and (4.4).

Next we recognize that (4.2) implies that the operator

$$C^{2m-2} + (-1)^{m-1}2\left(m\binom{2m}{m}\right)^{-1}(\Delta x)^{2m-2}D_+^{m-1}D_-^m, \tag{4.5}$$

is a $2m-1$ *order* approximation to $\frac{\partial}{\partial x}$, with stencil $(-(m-1),m)$. Moreover, the same is true for the operator

$$C^{2m} - \lambda_{m-1}^m(\Delta x)^{2m-1}D_+^mD_-^m = \tag{4.6}$$

$$= C^{2m} - (-1)^{m-1}\left(m\binom{2m}{m}\right)^{-1}(\Delta x)^{2m-1}D_+^mD_-^m$$

This operator is easily shown to be unique. Thus, we have the important result:

$$C^{2m} = C^{2m-2} + (-1)^{m-1}\left(m\binom{2m}{m}\right)^{-1}(\Delta x)^{2m-2}D_+^{m-1}D_-^m(2 + \Delta xD_+) \tag{4.7}$$

This translates to (using (4.4))

$$\sum_{k=-m+1}^{m-1} \nu_k^m\Delta_+\left(\Delta_+S^{k-1}\right) = \sum_{k=-m+2}^{m-2} \nu_k^{m-1}\Delta_+\left(\Delta_+S^{k-1}\right) \tag{4.8}$$

$$+ \sum_{k=1-m}^{m-2} 2\left(m\binom{2m}{m}\right)^{-1}(-1)^{k-1}\binom{2m-3}{m+k-1}\Delta_+\left(\Delta_+S^{k-1}\right)$$

$$+ \sum_{k=1-m}^{m-1} \left(m\binom{2m}{m}\right)^{-1}(-1)^{k}\binom{2m-2}{m+k-1}\Delta_+\left(\Delta_+S^{k-1}\right)$$

which implies

$$\nu_k^m = \nu_k^{m-1} + (-1)^k\left(m\binom{2m}{m}\right)^{-1}\binom{2m-2}{m+k-1}(m-1)^{-1}k \tag{4.9}$$

for $k=1,\ldots,m-1$. Thus we have proven (3.2) (again defining $\nu_m^m = 0$.)

It is now easy to see that

$$(-1)^k\nu_k^m > 0, \ \textit{for} \ \ k=1,\ldots,m-1. \tag{4.10}$$

Next we apply the identity (4.8) to the grid function defined by:

$$q_j = \frac{1-(-1)^j}{2} = -q_{-j} \ \textit{for} \ j \geq 0. \tag{4.11}$$

This leads us to the useful result:

$$\sum_{k=1}^{m-1}(-1)^k\nu_k^m = \sum_{k=1}^{m-2}(-1)^k\nu_k^{m-1} + \frac{1}{4}(2m-1), \qquad m = 2,3,\ldots \tag{4.12}$$

thus

$$\sum_{k=1}^{m-1}(-1)^k\ \nu_k^m = \frac{1}{4}\sum_{j=2}^{m}\frac{1}{(2j-1)}, \qquad m=2,3\ldots \tag{4.13}$$

(The fact that the series above diverges as $m \to \infty$ explains why $b_{\max} \to 0$ as $m \to \infty$.)

We may rewrite:

$$C^{2m} = D_+ + \sum_{k=-m+1}^{m-1} \mu_k^m D_+\left(\Delta_+S^{k-1}\right) \tag{4.14a}$$

This, together with (4.4a) gives us the identities:

$$\mu_o^m = -\frac{1}{2} \tag{4.14b}$$

$$\mu_k^m = \nu_k^m, k \neq O. \tag{4.14c}$$

Next, we claim that we can rewrite:

$$C^{2m} f(q_j) = \frac{1}{\Delta x} df^-_{j+\frac{1}{2}} + \sum_{k=-m+1}^{m-1} \mu_k^m D_+ \left(df^-_{j+k-\frac{1}{2}} \right) \tag{4.15}$$

$$+ \frac{1}{\Delta x} df^+_{j-\frac{1}{2}} + \sum_{k=-m+1}^{m-1} \nu_k^m D_- \left(df^+_{j+k-\frac{1}{2}} \right)$$

We verify this by rewriting the right side above as:

$$D_+ h(q_j, q_{j-1}) + \sum_{\substack{k=-m+1 \\ k \neq 0}}^{m-1} \nu_k^m D_+ (\Delta_- f(q_j))$$

$$+ \frac{1}{2} D_+ \Big[-h(q_j, q_{j-1}) + f(q_{j-1})$$

$$+ f(q_j) - h(q_j, q_{j-1}) \Big]$$

$$= D_0(\Delta x) f(q) + \sum_{\substack{k=-m+1 \\ k \neq 0}}^{m-i} \nu_k^m D_+ (\Delta_- f(q_j))$$

$$= D_+ S^{-1} f(q) + \sum_{k=-m+1}^{m-1} \nu_k^m D_+ (\Delta_- f(q_j)) = C^{2n} f(q_j)$$

We have thus rewritten the 2m*th* order, nondissipative approximation, C^{2m}, in terms of an arbitrary E flux, in a form convenient for the purpose of making it TVD.

Next we note that the approximation to

$$q_t = 0$$

of the form:

$$\frac{\partial q_j}{\partial t} = (-1)^{m-1}\beta(\Delta x)^{2m-1} D_+^m D_-^m q_j \tag{4.16}$$

$$= \beta \sum_{k=1-m}^{m-1} (-1)^k \binom{2m-2}{m+k-1} D_+(\Delta_+ S^{k-1}) q_j$$

is dissipative of order 2m, and accurate of order 2m-1. Its Fourier transform is easily seen to satisfy

$$\frac{\partial}{\partial t}\hat{q}(\zeta) = -\frac{\beta}{\Delta x}[2-2\cos(\zeta\,\Delta x)^m]\hat{q}(\zeta) \tag{4.17}$$

Thus, for an arbitrary E flux, we may write a $2m-1$ order scheme, with $2m-order$ dissipation, approximating (1.1) as:

$$\frac{\partial}{\partial t} q_j = -C^{2m} f(q_j) + (-1)^{m-1}\beta(\Delta x)^{2m-1} D_+^m D_-^{m-1}[df^+_{j-\frac{1}{2}} - df^-_{j-\frac{1}{2}}] \tag{4.18}$$

$$= -\left[\frac{df^+_{j+\frac{1}{2}}}{\Delta x} + \sum_{k=-m+1}^{m-1}\left[\mu_k^m + (-1)^k\beta\binom{2m-2}{m+k-1}\right] D_+\left(df^-_{j+k-\frac{1}{2}}\right)\right.$$

$$\left. + \frac{df^+_{j-\frac{1}{2}}}{\Delta x} + \sum_{k=-m+1}^{m-1}\left[\nu_k^m - (-1)^k\beta\binom{2m-2}{m+k-1}\right] D_+(df^+_{j+k-\frac{1}{2}})\right]$$

Thus, we have constructed the unlimited version of the numerical flux, $\hat{f}^{m,\beta}_{j+\frac{1}{2}}$, of (3.1), having the relevant desired properties.

For convenience, we require

$$\beta\binom{2m-2}{m+k-1} \leq |\mu_k^m| = |\nu_k^m| \quad \text{for } \textit{each} \ \ k = 0\ ,\pm 1 \ldots, +-(m-1) \tag{4.19}$$

We claim that this is true for all k, if it is true for $k = m-1$, or if

$$\beta \leq \left(m\binom{2m}{m}\right)^{-1} = |\nu_{m-1}^m|. \tag{4.20}$$

We shall prove this using induction and (4.9). The result is obviously always valid for $k = m-1$. Suppose (4.19) is valid for all $|k| \leq m-1$ and all numbers up to m. Then we have, from (4.9)

$$(-1)^k \nu_k^{m+1} - \left((m+1)\binom{2m+2}{m+1}\right)^{-1}\binom{2m}{m+k} \tag{4.21}$$

$$= (-1)^k \nu_k^m - \left(m\binom{2m}{m}\right)^{-1}\binom{2m-2}{m+k-1}$$

$$+ \left[\frac{k}{m}\binom{2m+2}{m+1}^{-1}(m+1)^{-1}\binom{2m}{m-k} + \left(m\binom{2m}{m}\right)^{-1}\binom{2m-2}{m+k-1}\right]$$

$$- (m+1)^{-1}\binom{2m+2}{m+1}^{-1}\binom{2m}{m+k}\Bigg]$$

We shall show that the last expression in brackets above is always positive.

For $k = 0$, we have:

$$\frac{1}{2\,(2m-1)} - \frac{1}{2\,(2m+1)} > 0$$

For $k = 1$, we have:

$$\frac{1}{2(m+1)(2m-1)} + \frac{m-1}{2m(2m-1)} + \frac{-m}{2(m+1)(2m+1)}$$

$$> (m-1)\left[\frac{1}{2m\,(2m-1)} - \frac{1}{2(m+1)(2m-1)}\right] > 0$$

For $k \geq 2$, we have:

$$\frac{1}{2m\,(2m-1)}\,\frac{(m-1)\ldots(m-k)}{(m+k-1)\cdots(m+1)}\left[\frac{1}{(m^2-k^2)}km\left(\frac{2m-1}{2m+1}\right)+1-\frac{m^2(2m-1)}{(m^2-k^2)(2m+1)}\right]$$

$$=\frac{1}{(2m)(2m-1)}\,\frac{(m-1)\ldots(m-k)}{(m+k-1)\ldots(m+1)}\left[\frac{2}{2m+1}+\left(\frac{2m-1}{2m+1}\right)\frac{k}{(m+k)}\right]>0$$

The claim is now proven.

Next we apply the flux limiter to (4.18), arriving at the scheme (2.2), (3.1)- (3.6). To verify that it decreases variation, we rewrite it as Equation (4.22):

$$\frac{\partial q_j}{\partial f}=D_+q_j\left[-\left(\frac{df^-_{j+\frac{1}{2}}}{\Delta_+q_j}\right)\left[1+\sum_{k=-m+1}^{m-1}\left(\mu^m_k+(-1)^k\beta\binom{2m-2}{m+k-1}\right)\left(\frac{\left(df^-_{j+k+\frac{1}{2}}\right)^{(k)}-\left(df^-_{j+k-\frac{1}{2}}\right)^{(k)}}{df^-_{j+\frac{1}{2}}}\right)\right]\right]$$

$$-D_-q_j\left[\left(\frac{df^+_{j-\frac{1}{2}}}{\Delta_-q_j}\right)\left[1+\sum_{k=-m+1}^{m-1}\left(\nu^m_k-(-1)^k\beta\binom{2m-2}{m+k-1}\right)\left(\frac{\left(df^+_{j+k-\frac{1}{2}}\right)^{(k)}-\left(df_{j+k-\frac{1}{2}}\right)^{(k)}}{df^+_{j-\frac{1}{2}}}\right)\right]\right]$$

$$=C_{j+\frac{1}{2}}D_+q_j-D_{j-\frac{1}{2}}D_-q_j$$

This is TVD and satisfies a maximum principle if, for each j:

$$C_{j+\frac{1}{2}},\;\; D_{j+\frac{1}{2}}\geq 0. \tag{4.23}$$

See e.g. [21].

Thus, we need:

$$1 \geq \sum_{k=-m+1}^{m-1} \left(\mu_k^m + (-1)^k \beta \binom{2m-2}{m+k-1}\right) \frac{\left[\left(df^-_{j+k-\frac{1}{2}}\right)^{(k)} - \left(df_{j+k+\frac{1}{2}}\right)^{(k)}\right]}{df^-_{j+\frac{1}{2}}} \tag{4.24a}$$

$$1 \geq \sum_{k=-m+1}^{m-1} \left(\nu_k^m - (-1)^k \beta \binom{2m-2}{m+k-1}\right) \frac{\left[\left(df^+_{j+k-\frac{1}{2}}\right)^{(k)} - \left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}\right]}{df^+_{j-\frac{1}{2}}} \tag{4.24b}$$

In (4.24b), we estimate the right side, using definition (3.5a, b, c), and recalling that the sign of the *kth* coefficient is $(-1)^k$ if $k \geq 0$, $(-1)^{k+1}$ if $k < 0$. Thus we need:

$$1 \geq \frac{1}{2} - \beta \binom{2m-2}{m-1} + \sum_{k=1}^{m-1} \left((-1)^k \nu_k^m - \beta \binom{2m-2}{m+k-1}\right) b \tag{4.25}$$

$$+ \sum_{k=-m+1}^{-1} \left((-1)^{k+1} \nu_k^m + \beta \binom{2m-2}{m+k-1}\right) b$$

or:

$$\frac{1}{2} + \beta \binom{2m-2}{m-1} \geq 2 \left(\sum_{k=1}^{m-1} (-1)^k \nu_k^m \right) b. \tag{4.26}$$

or (using (4.13)):

$$b \leq \frac{1 + 2\beta \binom{2m-2}{m-1}}{\sum_{k=2}^{m} \frac{1}{2k-1}} \leq \frac{1 + \frac{2}{m} \binom{2m}{m}^{-1} \binom{2m-2}{m-1}}{\sum_{k=2}^{m} \frac{1}{2k-1}}, \tag{4.27}$$

which implies $D_{j-\frac{1}{2}} \geq 0$.

A similar argument shows that $C_{j+\frac{1}{2}} \geq 0$ for the same values of b.

A simple exercise on a pocket calculator shows us that

$$\sum_{j=2}^{7} \frac{1}{2j-1} \approx .9551 \tag{4.28a}$$

$$\sum_{j=2}^{8} \frac{1}{2j-1} \approx 1.0218 \tag{4.28b}$$

$$\sum_{j=2}^{9} \frac{1}{2j-1} \approx 1.0809 \tag{4.28c}$$

It is possible to choose $\beta \leq \left(8\binom{16}{8}\right)^{-1}$ so that $b \geq 1$ because $\frac{1}{15} > .0218$. If it were possible to do this for $m = 9$, we would have $\frac{1}{17} > .0809$, which is false.

Thus, within our constraints, 15*th* order accuracy (in 17 points) is the highest possible.

Next, we obtain the CFL restriction for the explicit forward-Euler time discretization, which we write as

$$q_j^{n+1} - q_j^n = \frac{\Delta t}{\Delta x}\left[C^n_{j+\frac{1}{2}}\Delta_+ q_j^n - D_{j-\frac{1}{2}}\Delta_- q_j^n\right] \tag{4.29}$$

The precise restriction for the scheme to be TVD, in addition to (4.23) is, for each j:

$$\frac{\Delta t}{\Delta x}\left[C^n_{j+\frac{1}{2}} + D^n_{j+\frac{1}{2}}\right] \leq 1 \tag{4.30}$$

(see [13])

We thus wish to obtain upper bounds for

$$\left[1 + \sum_{k=-m+1}^{m-1}\left(\mu_k^m + (-1)^k\beta\binom{2m-2}{m+k-1}\right)\frac{\left[\left(df^-_{j+k+\frac{1}{2}}\right)^{(k)} - \left(df^-_{j+k-\frac{1}{2}}\right)^{(k)}\right]}{df^-_{j+\frac{1}{2}}}\right] \tag{4.31a}$$

and

$$\left[1 + \sum_{k=-m+1}^{m-1} \left(\nu_k^m - (-1)^k \beta \binom{2m-2}{m+k-1}\right) \frac{\left[\left(df^+_{j+k+\frac{3}{2}}\right)^{(k)} - \left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}\right]}{df^+_{j+\frac{1}{2}}}\right] \tag{4.31b}$$

A routine calculation using the definitions (3.5), noting the signs of the coefficients, gives us the result (d) in Theorem (3.1), modulo proving that

$$\nu_1^m = -\sum_{j=2}^{m} \frac{1}{2j(2j-1)}, \qquad m \geq 2. \tag{4.32}$$

For $m = 2$, (3.2a) gives:

$$\nu_1^2 = \left(2\binom{4}{2}\right)^{-1} = \frac{1}{12}.$$

Assume that (4.32) is valid up to m. Then, (4.9) for $k = 1$, gives us

$$\nu_1^{m+1} = \nu_1^m - \frac{1}{m+1}\binom{2m-2}{m+1}^{-1}\binom{2m}{m+1}\frac{1}{m}$$

$$= -\sum_{j=2}^{m}\frac{1}{2j(2j-1)} - \frac{1}{2(m+1)(2(m+1)-1)}$$

Finally, we check the stability of these linearized "β" schemes, without flux limiters, using explicit forward Euler time discretization. We linearize about a constant state $\bar{q}$, at which $f'(\bar{q}) = a \neq 0$.

This is:

$$\frac{q_j^{n+1} - q_j^n}{\Delta t} = -\left[C^{2m} a q_j^n + (-1)^{m-1}\frac{\beta}{\Delta x}\left(h_o(\bar{q}, \bar{q}) - h_1(\bar{q}, \bar{q})\right)(\Delta_+\Delta_-)^m q_j^n\right] \tag{4.33}$$

We note that h is an E flux. In [19] it was shown for such fluxes that:

$$h_o(\bar{q}, \bar{q}) \geq 0 \geq h_1(\bar{q}, \bar{q})$$

If equality holds for both above, then consistency implies

$$0 = h_o(\bar{q}, \bar{q}) + h_1(\bar{q}, \bar{q}) = f'(\bar{q}) = a \neq o.$$

which is a contradiction. Thus we may define the positive quantity

$$B = \beta(h_o(\bar{q}, \bar{q}) - h_1(\bar{q}, \bar{q})) > 0$$

The amplification matrix for (4.33) is

$$1 - \frac{a\Delta t}{\Delta x} i(\zeta + C(\zeta)\zeta^{2m+1}) - \frac{\Delta t}{\Delta x} B(2-2\cos\zeta)^m = A(\zeta) \tag{4.34}$$

for $-\pi \leq \zeta < \pi$, $C(0) \neq 0$, and $C(\zeta)$ real analytic for real ζ. Then the relation:

$$|A(\zeta)|^2 = 1 + a^2\left(\frac{\Delta t}{\Delta x}\right)^2 \zeta^2 - \frac{2\Delta t}{\Delta x} B\zeta^{2m} + 0(\zeta^{2m+1}) \leq 1. \tag{4.35}$$

which implies:

$$|a^2|\frac{\Delta t}{\Delta x} \leq 2\, B\zeta^{2m-2} + 0(\zeta^{2m-1}) \;\; as \;\; |\zeta|\downarrow 0.$$

This is a contradiction, since $m \geq 2$.

Theorem (3.1) is now proven.

To construct the "α" schemes of Theorem (3.2) we first construct a dissipative approximation to $q_t = 0$:

$$\frac{\partial q_j}{\partial t} = (-1)^m \, \alpha \, \Delta_+^{m-1} \Delta_-^{m-1} D_- q_j, \;\; for \;\; \alpha > 0 \tag{4.36}$$

The operator on the right above has module (-m,m-1), and its symbol is:

$$\alpha(-1)^m\, 2^{m-1}(\cos\zeta-1)^{m-1}(1-e^{-i\zeta}) \tag{4.37}$$

$$=\alpha 2^{m-1}[-(1-cos\zeta)^m - i(1-cos\zeta)^{m-1}\sin\zeta]$$

Thus, this operator is dissipative of order $2m$ and accurate of order $2m-2$. It may be rewritten as:

$$\frac{\partial q_j}{\partial t} = \alpha \sum_{k=-m+1}^{m-2} (-1)^k \binom{2m-3}{k+m-1} D_+(\Delta_+ S^{k-1} q_j) \tag{4.38}$$

Similarly, the operator on the right side of

$$\frac{\partial q_j}{\partial t} = (-1)^{m-1}\alpha\Delta_+^{m-1}\Delta_-^{m-1}D_+q_j \qquad \textit{for}\ \ \alpha > 0, \tag{4.39}$$

has module $(-m+1,m)$, and its symbol is:

$$\alpha 2^{m-1}[-(1-cos\zeta)^m + i(1-cos\zeta)^{m-1}\sin\zeta] \tag{4.40}$$

It is again dissipative of order $2m$, is $2m-2$ *order* accurate, and it may be rewritten as:

$$\frac{\partial q_j}{\partial t} = \alpha \sum_{k=-m+2}^{m-1} (-1)^k \binom{2m-3}{k+m-2} D_+(\Delta_+ S^{k-1} q_j). \tag{4.41}$$

Thus, we may use (4.15), (replacing m by $m-1$), (4.38) and (4.41) to obtain the unlimited "α" scheme.

$$\frac{\partial}{\partial t}q_j = -C^{2m-2}f(q_j) + (-1)^m\alpha(\Delta x)^{2m-2}\left[D_+^{m-1}D_-^{m-1}D_+h(q_j,q_{j-1})\right] \tag{4.42}$$

$$= -\left[\frac{1}{\Delta x}\, df^-_{j+\frac{1}{2}} + \sum_{k=-m+2}^{m-1}\left[\mu_k^{m-1} + (-1)^k\alpha\binom{2m-3}{k+m-2}\right]D_+\left(df^+_{j-k-\frac{1}{2}}\right)\right.$$

$$+ \frac{1}{\Delta x} df^{+}_{j-\frac{1}{2}} + \sum_{k=-m+1}^{m-2} \left[\nu_k^{m-1} - (-1)^k \alpha \binom{2m-3}{k+m-1} \right] D_+ \left(df^{+}_{j+k-\frac{1}{2}} \right) \Bigg]$$

Using (4.2), we see that the leading term of the truncation error of the right side of (4.42) is:

$$TE = (-1)^m \left[\alpha - 2 \frac{(m-1)!\, m!}{(2m)!} \right] (\Delta x)^{2m-2} \partial_x^{2m-1} f(q), \tag{4.43}$$

which is independent of the choice of the E flux, $h(q_{j+1}, q_j)$.

From (4.07), it is clear that α and β schemes coincide if $\alpha = \frac{2}{m} \binom{2m}{m}^{-1}$, and $\beta = \frac{1}{m} \left(\frac{2m}{m} \right)^{-1}$.

For convenience, we want (for $m \geq 2$):

$$\alpha \binom{2m-3}{k+m-2} \leq |\mu_k^{m-1}| = |\nu_k^{m-1}|, \; k = 0,1, \ldots\, m-2. \tag{4.44}$$

We claim that this is valid for all these k if it is valid for $k = m-2$, or if

$$\alpha \leq \left((m-1) \binom{2m-2}{m-1} \right)^{-1} \tag{4.45}$$

This is trivial for $m=2$. For $m > 2$, this reduces to showing that:

$$\left(m \binom{2m}{m} \right)^{-1} \binom{2m-1}{k+m-1} \leq |\nu_k^m|, \quad k = 0,1,\ldots,m-1. \tag{4.46}$$

We obtained a stronger inequality in (4.19), (4.20), so the validity of (4.44) from (4.45) is obvious.

Next we apply the flux limiters to (4.42), arriving at the scheme. (2.2), (3.7). To verify that the scheme decreases variation, we rewrite it as Equation (4.47):

$$\frac{\partial q_j}{\partial t} = D_+q_j\left[\frac{-df^-_{j+\frac{1}{2}}}{\Delta_+q_j}\left[1 + \sum_{k=-m+2}^{m-1}\left(\mu_k^{m-1} + (-1)^k\alpha\binom{2m-3}{m+k-1}\right)\left(\frac{\left(df^-_{j+k+\frac{1}{2}}\right)^{(k)} - \left(df^-_{j+k-\frac{1}{2}}\right)^{(k)}}{df^-_{j+\frac{1}{2}}}\right)\right]\right]$$

$$- D_-q_j\left[\frac{df^+_{j-\frac{1}{2}}}{\Delta_-q_j}\left[1+\sum_{k=-m+1}^{m-2}\left(\nu_k^{m-1} - (-1)^k\alpha\binom{2m-3}{m+k-1}\right)\left(\frac{\left(df^+_{j+k+\frac{1}{2}}\right)^{(k)} - \left(df^+_{j+k-\frac{1}{2}}\right)^{(k)}}{df^+_{j-\frac{1}{2}}}\right)\right]\right]$$

$$= C_{j+\frac{1}{2}}D_+q_j - D_{j-\frac{1}{2}}D_-q_j$$

We must show that (4.23) is valid for this scheme. Thus we need:

$$1 \geq \sum_{k=-m+2}^{m-1}\left(\mu_k^{m-1} + (-1)^k\alpha\binom{2m-3}{k+m-2}\right)\left(\frac{\left(df^-_{j+k-\frac{1}{2}}\right)^{(k)} - \left(df^-_{j+k+\frac{1}{2}}\right)^{(k)}}{df^+_{j+\frac{1}{2}}}\right) \tag{4.48a}$$

$$1 \geq \sum_{k=-m+1}^{m-2}\left(\nu_k^{m-1} - (-1)^k\alpha\binom{2m-3}{k+m-1}\right)\left(\frac{\left(df^+_{j+k-\frac{1}{2}}\right)^{(k)} - \left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}}{df^+_{j+\frac{1}{2}}}\right) \tag{4.48b}$$

In (4.48b), we estimate the right side, using the definition of ν_k^m from (4.4b), (4.4c), and (4.4d). We recall that the *kth* coefficient has sign $(-1)^k$ if $k \geq 0$, $(-1)^{k+1}$ if $k < 0$. Thus we need:

$$1 \geq \left[\frac{1}{2}-\alpha\binom{2m-3}{m-1}\right] \tag{4.49}$$

$$+\sum_{k=1}^{m-2}\left((-1)^k\nu_k^{m-1}-\alpha\binom{2m-3}{m+k-1}\right)b$$

$$+\sum_{k=-m+1}^{-1}\left((-1)^{k+1}\nu_k^{m-1}+\alpha\binom{2m-3}{m+k-2}\right)b \,.$$

or,

$$\frac{1}{2}+\alpha\binom{2m-3}{m-1} \geq 2\left(\sum_{k=1}^{m-2}(-1)^k\nu_k^{m-1}\right)b+\alpha\binom{2m-3}{m-2}b \tag{4.50}$$

Using (4.13) gives us:

$$b \leq \frac{1+2\alpha\binom{2m-3}{m-1}}{\sum_{k=2}^{m-1}\left(\frac{1}{2k-1}\right)+2\alpha\binom{2m-3}{m-2}} \tag{4.51}$$

The same inequality establishes (4.48a).

Using (4.28), we see that we can take $b > 1$ for $m \leq 8$, but not for $m = 9$. Thus 14*th* order (or 15*th* for $\alpha = \frac{2}{m}\binom{2m}{m}^{-1}$) in 17 points, is the best possible, as predicted.

Next we obtain the CFL restriction for the explicit time discretization (4.29), which requires inequality (4.30). This time it involves obtaining upper bounds for:

$$1+\sum_{k=-m+2}^{m-1}\left(\mu_k^{m-1}+(-1)^k\alpha\binom{2m-3}{m+k-2}\right)\left(\frac{\left(df^-_{j+k+\frac{1}{2}}\right)^{(k)}-\left(df_{j+k-\frac{1}{2}}\right)^{(k)}}{df^-_{j+\frac{1}{2}}}\right) \tag{4.52a}$$

and:

$$1+\sum_{k=-m+1}^{m-2}\left(\nu_k^{m-1}-(-1)^k\alpha\binom{2m-3}{m+k-1}\right)\left(\frac{\left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}-\left(df^+_{j-k-\frac{1}{2}}\right)^{(k)}}{df^+_{j-\frac{1}{2}}}\right) \tag{4.52b}$$

A routine calculation using the definitions in (3.5) and (4.32), and the signs of each coefficient, gives us result (e) in the statement of Theorem (3.2).

Finally, we check the stability of these linearized "α" schemes. We again linearize about a constant state $\bar{q}$, at which $f'(\bar{q}) = a \neq o$. The resulting scheme is as follows in Equation (4.53):

$$\frac{q_j^{n+1} - q_j^n}{\Delta t} = -\left[C^{2m-2} a\ q_j + (-1)^{m-1} \alpha (\Delta x)^{2m-2} D_+^{m-1} D_-^{m-1} D_+ [h_1(\bar{q},\bar{q}) q_j + h_o(\bar{q},\bar{q}) q_{j-1}\right]$$

We also know that:

$$h_0(\bar{q},\bar{q}) \geq 0 \geq h_1(\bar{q},\bar{q})$$

with at least one of these inequalities being strict.

The amplification matrix for (4.53) is:

$$1 - \frac{\Delta t}{\Delta x} ai[\zeta + C(\zeta)\ \zeta^{2m-1}] \tag{4.54}$$

$$- \frac{\Delta t}{\Delta x} \alpha\ 2^{m-1} (1 - cos\ \zeta)^{m-1} [h_o(\bar{q},\bar{q})[1 - cos\ \zeta + i \sin \zeta]$$

$$- h_1(\overline{q},\bar{q})[1 - \cos \zeta - i \sin \zeta]]$$

The rest of the proof goes as in (4.35).

Theorem (3.2) is now proven.

V. A More General Class of TVD Schemes

Given a conservation form approximation to the scalar version of (1.1) of the type:

$$q_j^{n+1} = q_j^n - \lambda\left(\hat{f}^n_{j+\frac{1}{2}} - \hat{f}^n_{j-\frac{1}{2}}\right) \tag{5.1}$$

where

$$\hat{f}^n_{j+\frac{1}{2}} = h(q^n_{j+k}, \ldots, q^n_{j-k+1})$$

Suppose we can rewrite

$$\hat{f}^n_{j+\frac{1}{2}} - \hat{f}^n_{j-\frac{1}{2}} = -\sum_{\nu=-k}^{k-1} A^{(\nu)}_{j+\frac{1}{2}}\, \Delta_+ q^n_{j+\nu} \tag{5.2}$$

subject to the following restrictions for each j:

$$A^{(k-1)}_{j+\frac{1}{2}} \geq 0 \geq A^{(-k)}_{j+\frac{1}{2}} \tag{5.3a}$$

$$A^{(\nu-1)}_{j+\frac{1}{2}} \geq A^{(\nu)}_{j-\frac{1}{2}} \quad \textit{for} \quad -k+1 \leq \nu \leq k-1,\ \nu \neq 0 \tag{5.3b}$$

$$1 \geq \lambda\left(A^{(0)}_{j-\frac{1}{2}} - A^{(0)}_{j+\frac{1}{2}}\right), \quad \textit{(the CFL restriction)} \tag{5.3c}$$

Then we have the following:

Theorem (5.1)

Given an approximation to (1.1), of the form (5.1), satisfying (5.2),(5.3), then the scheme is TVD, i.e

$$\sum_j |\Delta_+ q_j^{n+1}| \leq \sum_j |\Delta_+ q_j^n|$$

Proof:

Using a, by now, standard argument -e.g. [1], [13] and [28], we first compute:

$$(5.4) \quad \Delta_+ q_j^{n+1} = \lambda\, A^{(k-1)}_{j+\frac{3}{2}}\, \Delta_+ q^n_{j+k} - \lambda\, A^{(-k)}_{j+\frac{1}{2}}\, \Delta_+ q^n_{j-k}$$

$$+ \left(1 + \lambda\left(A^{(-1)}_{j+\frac{3}{2}} - A^{(0)}_{j+\frac{1}{2}}\right)\right) \Delta_+ q^n_j$$

$$+ \lambda \sum_{\substack{\nu=-k+1 \\ \nu\neq 0}}^{k-1} \left(-A^{(\nu)}_{j+\frac{1}{2}} + A^{(\nu-1)}_{j+\frac{3}{2}}\right) \Delta_+ q^n_{j+\nu}$$

Inequalities (5.3a, b, and c) were chosen so that each coefficient of $\Delta_+ q^n_{j+\nu}$ in (5.4) is non-negative. Thus we may take the absolute value of both sides, obtaining the inequality:

$$(5.5) \quad |\Delta_+ q_j^{n+1}| \leq |\Delta_+ q^n_j| + \lambda\, A^{(k-1)}_{j+\frac{3}{2}}\, |\Delta_+ q^n_{j+k}\,| - \lambda\, A^{(-k)}_{j+\frac{1}{2}}\, |\Delta_+ q^n_{j-k}\,|$$

$$+ \lambda \sum_{\nu=-k+1}^{k-1} \left(-A^{(\nu)}_{j+\frac{1}{2}} + A^{(\nu-1)}_{j+\frac{3}{2}}\right) |\Delta_+ q^n_{j+\nu}|$$

$$= |\Delta_+ q^n_j| + \lambda \Delta_+ \sum_{\nu=-k}^{k-1} A^{(\nu)}_{j+\frac{1}{2}} |\Delta_+ q^n_{j+\nu}|$$

We sum the inequality (5.5) over j, the result follows.

Next we approximate (1.1) via a semi-discrete method (2.2), where

$$(5.6) \qquad \hat{f}_{j+\frac{1}{2}} - \hat{f}_{j-\frac{1}{2}} = -\sum_{\nu=-k}^{k-1} A^{(\nu)}_{j+\frac{1}{2}}\, \Delta_+ q_{j+\nu}$$

Here the $A^{(\nu)}_{j+\frac{1}{2}}$ satisfy (5.3a and b). Then we have the following:

Theorem (5.2)

Given an approximation to (1.1) of the form (2.2), satisfying (5.6) and (5.3a and b), then the scheme is TVD, i.e.:

$$\frac{\partial}{\partial t} \sum_j |\Delta_+ q_j| \le 0$$

Proof

We follow an idea of Sanders [25], used by us in [21]. Let

$$\chi_{j+\frac{1}{2}}(t) = sgn\ \Delta_+ q_j$$

Then

$$\frac{\partial}{\partial t}|\Delta_+ q_j| = \chi_{j+\frac{1}{2}} \frac{\partial}{\partial t}\Delta_+ q_j \tag{5.7}$$

$$= A^{(k-1)}_{j+\frac{3}{2}}\, \chi_{j+\frac{1}{2}}\, \Delta_+ q_{j+k} - A^{(-k)}_{j+\frac{1}{2}}\, \chi_{j+\frac{1}{2}}\, \Delta_+ q_{j-k}$$

$$+ \sum_{\substack{\nu=-k+1 \\ \nu\neq 0}}^{k-1} \left[-A^{(\nu)}_{j+\frac{1}{2}} + A^{(\nu-1)}_{j+\frac{3}{2}} \right] \chi_{j+\frac{1}{2}}\, \Delta_+ q_{j+\nu}$$

$$+ |\Delta_+ q_j| \left[-A^{(0)}_{j+\frac{1}{2}} + A^{(-1)}_{j+\frac{3}{2}} \right]$$

Because of (5.3a and b), all the coefficients of $\chi_{j+\frac{1}{2}}\, \Delta_+ q_{j+\nu}$ for $\nu \neq 0$, are non-negative.

Thus, we have

$$\frac{\partial}{\partial t} |\Delta_+ q_j| \le \sum_{\nu=-k+1}^{k-1} \left[-A^{(\nu)}_{j+\frac{1}{2}} + A^{(\nu-1)}_{j+\frac{3}{2}} \right] |\Delta_+ q_{j+\nu}| \tag{5.8}$$

$$+A_{j+\frac{3}{2}}^{(k-1)} \, |\Delta_+ q_{j+k}| \, - A_{j+\frac{1}{2}}^{(-k)} \, |\Delta_+ q_{j-k}|$$

$$= \Delta_+ \sum_{\nu=-k}^{k-1} A_{j+\frac{1}{2}}^{\nu} \, |\Delta_+ q_{j+\nu}|$$

Summing this inequality gives us Theorem (5.2).

We hope that this very general approach to the construction of TVD schemes will lead to an even wider class of useful high order accurate methods. We shall discuss this in a future paper.

VI. Extensions to Hyperbolic Systems of Conservation Laws

We shall approximate such a system (1.1), using the scalar TVD approximations developed in sections II and III. The key tool in this construction is the use of a nonlinear field-by-field decomposition, which effectively decouples the system. Several such decompositions exist - Godunov's [9], Osher's [22], and Roe's [23]. For simplicity of exposition, we shall only use the last here.

Although the formal version of the last method is well known to violate the entropy condition -i.e. to have stable expansion shocks, it is possible to remove this difficulty, e.g. [19], by changing the differencing slightly near sonic points - points where $\lambda_k(q) = 0$ for some eigenvalue k. What we shall do here can be viewed as an extension of some of the work in [13] and [21] to higher order, non-oscillatory methods. For simplicity of exposition only, we shall ignore the entropy difficulty.

Given two states q_j, q_{j+1}, Roe [23] defines a matrix $A_{j+\frac{1}{2}}$ satisfying the equality:

$$f(q_{j+1}) - f(q_j) = A_{j+\frac{1}{2}}(q_{j+1} - q_j) \tag{6.1}$$

This matrix is supposed to depend continuously on q_j, q_{j+1}, to have only real eigenvalues, and to satisfy

$$\lim_{q_{j+1} \to q_j} A_{j+\frac{1}{2}} = \partial f(q_j)$$

Such a matrix exists if a convex entropy exists for the system [11]. See [23],[24] for some special properties of physical systems.

Let the eigenvalues of $A_{j+\frac{1}{2}}$ be denoted by $\lambda^{(p)}_{j+\frac{1}{2}}$, $p = 1,\ldots,n$. The corresponding left eigenvectors are $l^{(p)}_{j+\frac{1}{2}}$, and right eigenvectors are $r^{(p)}_{j+\frac{1}{2}}$, normalized so that

$$l^{(p)}_{j+\frac{1}{2}} \cdot r^{(q)}_{j+\frac{1}{2}} = \delta_{pq} = 1 \text{ if } p = q$$

$$= 0 \text{ if } p \neq q$$

Then we may write

$$q_{j+1} - q_j = \sum_{p=1}^{n} \alpha^{(p)}_{j+\frac{1}{2}} r^{(p)}_{j+\frac{1}{2}} \tag{6.2a}$$

$$f(q_{j+1}) - f(q_j) = \sum_{p=1}^{n} \lambda^{(p)}_{j+\frac{1}{2}} \alpha^{(p)}_{j+\frac{1}{2}} r^{(p)}_{j+\frac{1}{2}} \tag{6.2b}$$

Let

$$x^+ = \max(x,0)$$
$$x^- = \min(x,0) \tag{6.3}$$

Next we define Roe's first order numerical flux:

$$h(q_{j+1},q_j) = \frac{1}{2}(f(q_{j+1} + f(q_j)) - \frac{1}{2}|A_{j+\frac{1}{2}}|(q_{j+1}-q_j) \tag{6.4}$$

so

$$df^-_{j+\frac{1}{2}} = \sum_{p=1}^{n} (\lambda^{(p)}_{j+\frac{1}{2}})^- \alpha^{(p)}_{j+\frac{1}{2}} r^{(p)}_{j+\frac{1}{2}} \tag{6.5a}$$

$$df^+_{j+\frac{1}{2}} = \sum_{p=1}^{n} (\lambda^{(p)}_{j+\frac{1}{2}})^+ \alpha^{(p)}_{j+\frac{1}{2}} r^{(p)}_{j+\frac{1}{2}} \tag{6.5b}$$

Now we use the notation of section III to construct the high order non-oscillatory scheme for systems of conservation laws. Let the quantities ν^m_k, μ^m_k, b, β, and α, be defined as in those sections. The numerical fluxes used to construct semi-discrete approximations of the form (1.7a) are defined via:

(β *schemes of* $2m-1$ *order accuracy*):

$$\hat{f}_{j+\frac{1}{2}} = \hat{f}^{m,\beta}_{j+\frac{1}{2}} = h(q_{j+1},q_j) \tag{6.6}$$

$$+ \sum_{k=-m+1}^{m-1} \left(\mu_k^m + (-1)^k \beta \binom{2m-2}{k+m-1}\right) \left(df^-_{j+k+\frac{1}{2}}\right)^{(k)}$$

$$+ \sum_{k=-m+1}^{m-1} \left(\nu_k^m - (-1)^k \beta \binom{2m-2}{k+m-1}\right) \left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}$$

or

(α schemes of $2m-2$ *or* $2m-1$ order accuracy)

$$\hat{f}_{j+\frac{1}{2}} = \hat{f}^{m,\alpha}_{j+\frac{1}{2}} = h(q_{j+1},q_j) \tag{6.7}$$

$$+ \sum_{k=-m+2}^{m-1} \left(\mu_k^{m-1} + (-1)^k \alpha \binom{2m-3}{k+m-2}\right) \left(df^-_{j+k+\frac{1}{2}}\right)^{(k)}$$

$$+ \sum_{k=-m+1}^{m-2} \left(\nu_k^{m-1} - (-1)^k \alpha \binom{2m-3}{k+m-1}\right) \left(df^+_{j+k+\frac{1}{2}}\right)^{(k)}$$

Now we define these vector valued flux limited quantities as follows:

$$[df^-_{j+\frac{1}{2}}]^{(k)} = \sum_{p=1}^{n} \min\bmod\,[(\lambda^{(p)}_{j+\frac{1}{2}})^- \alpha^{(p)}_{j+\frac{1}{2}},\, b(\lambda^{(p)}_{j-k+\frac{1}{2}})^- \alpha^{(p)}_{j-k+\frac{1}{2}},\, b(\lambda^{(p)}_{j-k+\frac{3}{2}})^- \alpha^{(p)}_{j-k+\frac{3}{2}})] r^{(p)}_{j+\frac{1}{2}}, \tag{6.8a}$$

for all k with $0 \neq k \neq 1$.

$$[df^-_{j+\frac{1}{2}}]^{(0)} = \sum_{p=1}^{n} \min\bmod\,[(\lambda^{(p)}_{j+\frac{1}{2}})^- \alpha^{(p)}_{j+\frac{1}{2}},\, b(\lambda^{(p)}_{j+\frac{3}{2}})^- \alpha^{(p)}_{j+\frac{3}{2}}]\, r^{(p)}_{j+\frac{1}{2}} \tag{6.8b}$$

$$[df^{-}_{j+\frac{1}{2}}]^{(1)} = \sum_{p=1}^{n} \min\bmod\,[(\lambda^{(p)}_{j+\frac{1}{2}})^{-}\,\alpha^{(p)}_{j+\frac{1}{2}},\, b(\lambda^{(p)}_{j-\frac{1}{2}})^{-}\,\alpha^{(p)}_{j-\frac{1}{2}}]\, r^{(p)}_{j+\frac{1}{2}} \tag{6.8c}$$

$$[df^{+}_{j+\frac{1}{2}}]^{(k)} = \sum_{p=1}^{n} \min\bmod\,[(\lambda^{(p)}_{j+\frac{1}{2}})^{+}\,\alpha^{(p)}_{j+\frac{1}{2}},\, b(\lambda^{(p)}_{j-k+\frac{1}{2}})^{+}\,\alpha^{(p)}_{j-k+\frac{1}{2}},\, b(\lambda^{(p)}_{j-k-\frac{1}{2}})^{-}\,\alpha^{(p)}_{j-k-\frac{1}{2}}]\, r^{(p)}_{j+\frac{1}{2}} \tag{6.8d}$$

for all k with $0 \neq k \neq -1$

$$[df^{+}_{j+\frac{1}{2}}]^{(0)} = \sum_{p=1}^{n} \min\bmod\,[(\lambda^{(p)}_{j+\frac{1}{2}})^{+}\,\alpha^{(p)}_{j+\frac{1}{2}},b\,(\lambda^{(p)}_{j-\frac{1}{2}})^{+}\,\alpha^{(p)}_{j-\frac{1}{2}}]\, r^{(p)}_{j+\frac{1}{2}} \tag{6.8e}$$

$$[df^{+}_{j+\frac{1}{2}}]^{(-1)} = \sum_{p=1}^{n} \min\bmod\,[(\lambda^{(p)}_{j+\frac{1}{2}})^{+}\,\alpha^{(p)}_{j+\frac{1}{2}},\, b(\lambda^{(p)}_{j+\frac{3}{2}})^{+}\,\alpha^{(p)}_{j+\frac{1}{2}}]\, r^{(p)}_{j+\frac{1}{2}} \tag{6.8f}$$

It is easily seen that each of the unlimited versions of these semi-discrete algorithms does indeed have the desired accuracy, for general nonlinear systems of hyperbolic conservation laws.

In the special case of linear diagonalizable hyperbolic systems:

$$f(q) = A\,q = A_{j+\frac{1}{2}}q, \text{ for each } j+\frac{1}{2},$$

we have a great deal of theory. We may now use the $\alpha^{(p)}_{j+\frac{1}{2}}$ to help measure the variation.

Define

$$|\Delta_{+}q_j| = \sum_{p=1}^{n} |\alpha^{(p)}_{j+\frac{1}{2}}| \tag{6.9}$$

A scheme of this semi-discrete type is said to be TVD if

$$\frac{d}{dt}\sum_{j} |\Delta_{+}q_j| \leq 0$$

Also the ratio needed below is defined by its value on the right:

$$\left| \frac{df^{+}_{j+\frac{1}{2}} - df^{-}_{j+\frac{1}{2}}}{\Delta_{+} q_j} \right| = \sup_p \left| \lambda^{(p)}_{j+\frac{1}{2}} \right| \tag{6.10}$$

We now have:

Theorem (6.1)

All the results of Theorems (3.1) and (3.2) go over word for word to the corresponding schemes for systems where the flux is defined by (6.4) to (6.8), and where $f(q) = Aq = A_{j+\frac{1}{2}} q$.

No theory of this type is known for nonlinear systems. The entropy condition was proven for bounded a.e. limits of a special second order TVD type approximation, using Osher's flux-decomposition in [21]. We find numerically that these schemes work quite well for compressible inviscid gas dynamical flows at widely varying Mach numbers. See [5] for the results of several numerical experiments.

VII. Results of Numerical Experiments

We now discuss some numerical results. Some members of the new family of schemes were programmed for a linear wave equation with a source term which drives the solution to a time-asymptotic steady state.

$$q_t + q_x - \pi\, cos(\pi x) = 0. \tag{7.1}$$

The semi-discrete TVD spatial differencing was combined with a family of multi-stage time differencing (which includes the simple one-stage scheme shown in Eq. (2.8)). The steady state exact solution of Eq. 7.1 is given by

$$q(x) = sin(\pi x) \tag{7.2}$$

The l_2 norm of the difference between the numerical and analytic steady-state solutions was computed and is presented below in Table (7.1) for a first-order accurate TVD scheme, and for the TVD and unlimited forms of some members of the new family of schemes.

We now discuss the results shown in Table (7.1). The last column entitled "global accuracy" is the order of accuracy measured from the numerical results. The order of accuracy is first measured based on the 20 interval and 30 interval solutions. Then, it is measured based on the 30 interval and the 40 interval solutions. Lastly, it is measured based on the 20 interval and the 40 interval solutions. The average of these three values have been entered in the last column of Table (7.1). The individual orders of accuracy (for every pair of intervals) is computed as follows: let the l_2 norm of the error for J intervals be denoted by E_J; then, the order of global (corresponding to the overall solution) error, O, is given by

$$0_{J1,J2} = \frac{ln(E_{J2}) - ln(E_{J1})}{ln(J2) - ln(J1)} \tag{7.3}$$

where $J1$ and $J2$ denote the number of intervals in the pair of solutions being considered.

Many facts stand out clearly in an analysis of the results of the 5-point schemes. The *TE* of the unlimited schemes is close to the theoretically derived values. The global accuracy of the TVD schemes shows some variation. In fact, the global accuracy of the TVD schemes based on Fromm's discretization and the Low *TE* Second-Order discretization compare quite favorably with the global accuracy of the Third-Order TVD scheme. In the case of the first two, the TVD scheme has better accuracy than the corresponding unlimited scheme. In the case of the Third-Order scheme, the accuracy suffers by going to the TVD form. When we consider the magnitude of error as

	l_1 norm of Error			
Scheme	20 intervals	30 intervals	40 intervals	Global Accuracy;
First-Order Accurate Monotone Upwind Scheme;				
First-Order	0.1496	0.1013	0.07662	$(\Delta x)^{0.97}$;
Third-Order Scheme, $\alpha = 1/6, b = 4$;				
Unlimited	0.0024856	0.000744	0.0003162	$(\Delta x)^{2.97}$;
TVD Limited	0.004212	0.001348	0.00076338	$(\Delta x)^{2.42}$;
Fully Upwind Scheme, $\alpha = \frac{1}{2}$ $b = 2$;				
Unlimited	0.019717	0.0089566	0.005091	$(\Delta x)^{1.95}$;
TVD Limited	0.017874	0.00784	0.004756	$(\Delta x)^{1.89}$;
Fromm's Scheme, $\alpha = \frac{1}{4}$, $b = 3$;				
Unlimited	0.005685	0.00239	0.0013221	$(\Delta x)^{2.10}$;
TVD Limited	0.00862	0.002773	0.001554	$(\Delta x)^{2.43}$;
Low *TE* Second-Order Scheme, $\alpha = \frac{1}{8}$, $b = 5$;				
Unlimited	0.003125	0.001256	0.000681	$(\Delta x)^{2.19}$;
TVD Limited	0.006767	0.0014528	0.001058	$(\Delta x)^{2.52}$;
TVD Central Difference Scheme, $\alpha = 0$ $b >> 1$;				
Smoothed	0.0100006	0.0045107	0.002556	$(\Delta x)^{1.97}$;
TVD Limited	0.02655	0.00559	0.0080886	$(\Delta x)^{1.275}$;
Unnamed Scheme, $\alpha = \frac{1}{3}$, $b = 5/2$				
Unlimited	0.01028809	0.004565	0.0025733	$(\Delta x)^{2.00}$;
TVD Limited	0.0111646	0.0045198	0.00268	$(\Delta x)^{2.09}$;

Table (7.1) Error Computations for some of the New 2nd and 3rd Order TVD Schemes opposed to the order of accuracy, the Third-Order scheme comes out ahead of all the others. The global order of accuracy of a TVD scheme depends on a number of factors, such as the number of maxima and minima, the ratio of this number to the overall number of intervals, implementation of boundary conditions, etc. Thus, the global accuracy of the TVD and the unlimited forms can be different. On the other hand, the fact that the Third-Order scheme is indeed third-order accurate in its unlimited form and that it consistently has a lower magnitude of error seems to imply that the Third-Order scheme may be the most preferable of the lot. The other second-order schemes having a low truncation error also suggest themselves as schemes which must be given serious consideration. We do not recommend the use of the unlimited forms of the TVD schemes whether the order of accuracy of these is higher or lower than the corresponding TVD formulation. The errors of the unlimited forms are shown here only for comparison. The TVD Central scheme is also highly unreliable as shown by the fact that its error for 30 intervals was actually better than the error of 40 intervals. This is due to the lack of dissipation. It has already been mentioned that the orders of global accuracy given in the table are the average of three values. It is quite instructive to actually look at the individual values that are averaged. Some schemes show a wider variation than others. The last remark here is that the Fully Upwind scheme, that many researchers (including the present authors) have been using in the recent past, is just about the worst of the lot (excluding the highly unreliable TVD Central scheme). In fact, to obtain the same level of accuracy as the 20 interval solution using the Third-Order scheme, the Fully Upwind scheme would need to use 40 intervals. The purely centrally differenced scheme shown in Table (7.1) as the Smoothed Central scheme (non-TVD central differencing along with a very small amount of third-order fourth-difference smoothing) does not fare much better when compared with the other third-order and second-order accurate schemes. The fifth order accurate, seven-point scheme leads to the following results:

l_1 norm of Error

Fifth Order, $\beta = \frac{1}{60} = \beta_{max}, b = \frac{9}{4}$

	20 Intervals	30 Intervals	40 Intervals	80 Intervals;
unlimited	.0000489	.00000662	.000001181;	
TVD limited	.0148	.00142	.0014	.000168;
(TVD limited)*	.0104	.00046	.00056;	

Table (7.2) Error Computation for 5th Order TVD Scheme

Here the (TVD limited)* line denotes calculating the l_1 norm of the error computed only at points where limiting does not occur - i.e., at which the scheme is of minimal order of accuracy. This measures the effect of pollution into high-accuracy regions.

Next in Figures (7.1a-e) we test the compressive properties of various approximations. We solve $q_t = -q_x$ with an initial Heaviside function. The third-order accurate, Fromm's, and the low error second-order scheme appear to be extremely accurate - as accurate as the scheme favored by Sweby in [26]. The fully upwind scheme has more smearing, while the first order accurate upwind method smears the profile excessively.

Bibliography

[1] L. Abrahamsson and S. Osher, Monotone difference schemes for singular pertubation problems, SIAM J. Num. Anal., V. 19 (1982), pp. 979-991.

[2] S. R. Chakravarthy, Relaxation Methods for Unfactored Implicit Upwind Schemes, AIAA-84-0165, Reno, NA (1984)

[3] S. R. Chakravarthy and S. Osher, High resolution applications of the Osher upwind scheme for the Euler equations, Proc. AIAA Comp. Fluid Dynamics Conf., Danvers, Mass (1983), pp. 363-372.

[4] S. R. Chakravarthy and S. Osher, Computing with High Resolution Upwind Schemes for Hyperbolic equations, to appear in Proceedings of AMS-SIAM, 1983 Summer Seminar, La Jolla, CA.

[5] S. R. Chakravarthy and S. Osher, A new class of High Accuracy Total Variation Diminishing Schemes for Hyperbolic Conservation Laws, In Preparation.

[6] S. R. Chakravarthy, K. Y. Szema, S. Osher, J. Gorski, A new class off High Accuracy Total Variation Diminishing Schemes for the Navier-Stokes Equations, In Preparation.

[7] P. Colella and P. R. Woodward, The piecewise-parabolic method (PPM) for gas-dynamical simulations, LBL report #14661, (July 1982).

[8] R. J. DiPerna, Convergence of approximate solutions to conservation laws, Arch., Rat. Mech. and Analysis, 82 (1983), pp. 27-70.

[9] S. K. Godunov, A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics, Mat. Sb., 47 (1959). pp. 271-290.

[10] J. B. Goodman and R. J. LeVeque, On the accuracy of stable schemes for two dimensional conservation laws, Math. Comp., (to appear).

[11] A. Harten, On the Symmetric Form of Systems of Conservation Laws with Entropy, ICASE Rep. No. 81-34, (1981), NASA Langley Research Center, Va.

[12] A. Harten, On a class of High Resolution Total-Variation-Stable Finite- Difference Schemes, SINUM, v.21, pp. 1-23 (1984).

[13] A. Harten, High resolution schemes for hyperbolic conservation laws, J. Comp. Phys., 49(1983), pp. 357-393.

[14] H. O. Kreiss and J. Oliger, Methcds for the Approximate Solution of Time Dependent Problems, GARP Publication series No. 10, (1973).

[15] S. N. Kruzkov, First order quasi-linear equations in several independent variables, Math. USSR Sb., 10 (1970), pp. 217-243.

[16] A. Majda and S. Osher, Numerical viscosity and the entropy condition, Comm. Pure Appl. Math., V. 32 (1979), pp. 797-838.

[17] W. A. Mulder and B. Van Leer, Implicit upwind computations for the Euler equations, AIAA Comp. Fluid Dynamics Conf., Danvers, Mass., (1983), pp. 303- 310.

[18] S. Osher, Numerical solution of singular perturbation problems and hyperbolic systems of conservation laws, North Holland Mathematical Studies #47, eds. S. Axelsson, L. S. Frank, and A. van der Sluis, pp. 179-205.

[19] S. Osher, Riemann solvers, the entropy condition, and difference approximations, SINUM, v. 21, (1984), pp. 217-235.

[20] S. Osher, Convergence of Generalized MUSCL Schemes, NASA Langley Contractor Report 172306, (1984), Submitted to SINUM.

[21] S. Osher and S. R. Chakravarthy, High resolution schemes and the entropy condition, SINUM, (to appear).

[22] S. Osher and F. Solomon, Upwind schemes for hyperbolic systems of conservation laws, Math. Comp., V. 38 (1982), pp. 339-377.

[23] P. L. Roe, Approximate Riemann solvers, parameter vectors, and difference schemes, J. Comp. Phys., V. 43 (1981), pp. 357-372.

[24] P. L. Roe, Some contributions to the modelling of discontinuous flows, to appear in Proceedings of AMS-SIAM 1983 Summer Seminar, La Jolla, CA.

[25] R. Sanders, On convergence of monotone finite difference schemes with variable spatial differencing, Math. Comp., v.40 (1983), pp. 91-106.

[26] P. K. Sweby, High resolution schemes using flux limiters for hyperbolic conservation laws, SINUM (to appear).

[27] E. Tadmor, Numerical viscosity and the entropy condition for conservative difference schemes, NASA Contractor Report 172141, (1983), NASA Langley, Math Comp. (to appear).

[28] B. Van Leer, Towards the ultimate conservative scheme, II. Monotonicity and conservation combined in a second order scheme, J. Comp. Phys. 14 (1974), pp. 361-376.

[29] B. Van Leer, Towards the Ultimate Conservative Finite Difference Scheme III. Upstream-Centered Finite-Difference Schemes for Ideal Compressible Flow, J. Comp. Phys., v. 23, (1977), pp. 263-275.

[30] B. Van Leer, Towards the ultimate conservative difference scheme. IV. A New apprcach to numerical convection, J. Comp. Phys., 23 (1977), pp. 276-298,

[31] H. C. Yee, R. F. Warming, and A. Harten, Implicit total variation diminishing (TVD) shemes for steady state calculations, Proc. AIAA Comp. Fluid Dynamics Conf., Danvers, Mass., (1983), pp. 110-127.

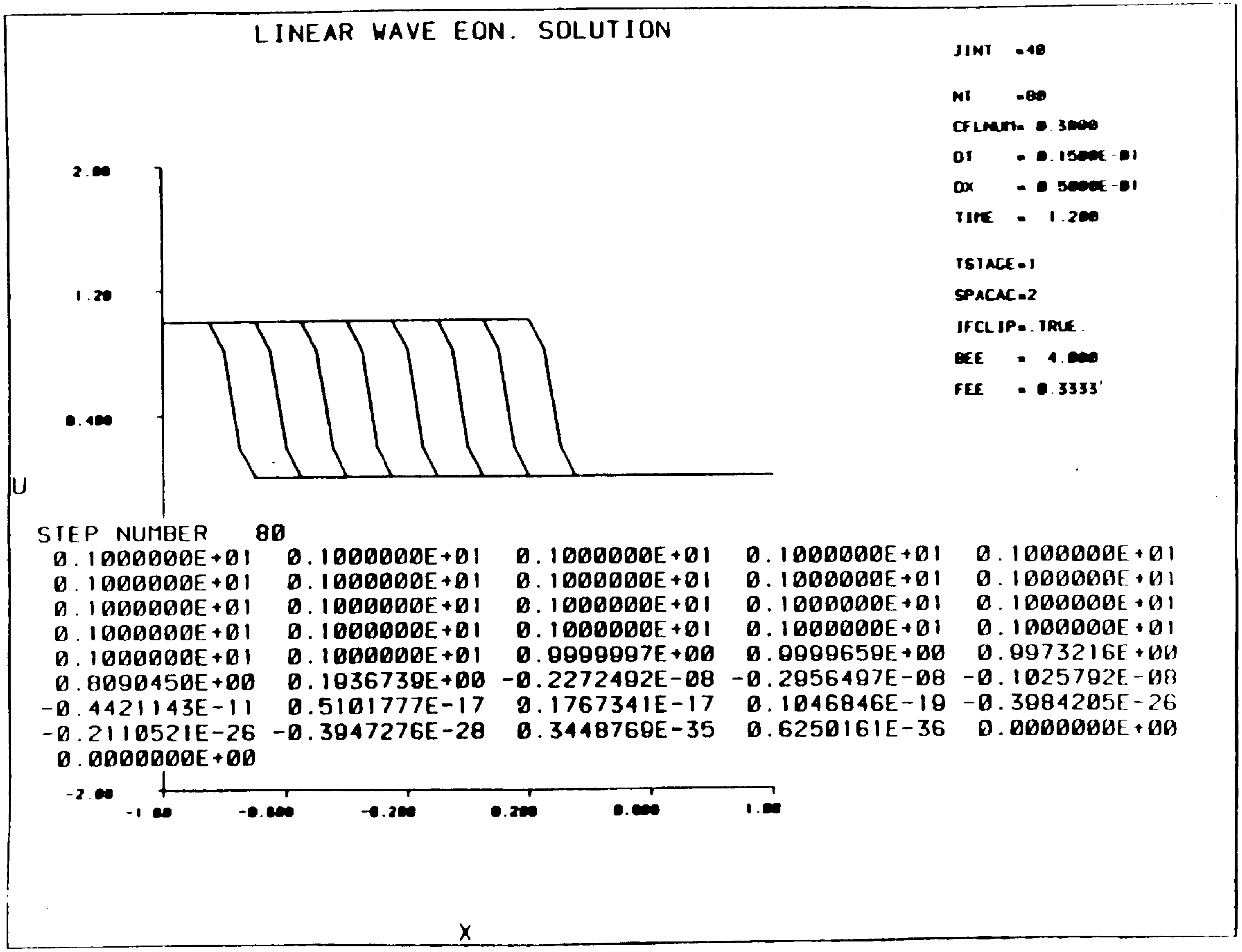
[32] S. T. Zalesak, Fully Multidimensional Flux-Corrected Transport, J. Comp. Phys., v. 31, (1979), pp. 335-362.

[33] B. Engquist and S. Osher, Stable and entropy condition satisfying approximations for transonic flow calculations. Math. Comp., 34 (1980), pp. 45-75.

[34] J. P. Boris and D. L. Book, Flux-Corrected Transport I - SHASTA, A fluid transport algorithm that works, J. Comp. Phys., v. 11, (1973), pp. 38- 69.

[35] P. D. Lax, Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves, SIAM Regional Conference Lectures in Applied Mathematics No. 11 (1972)

THIRD ORDER



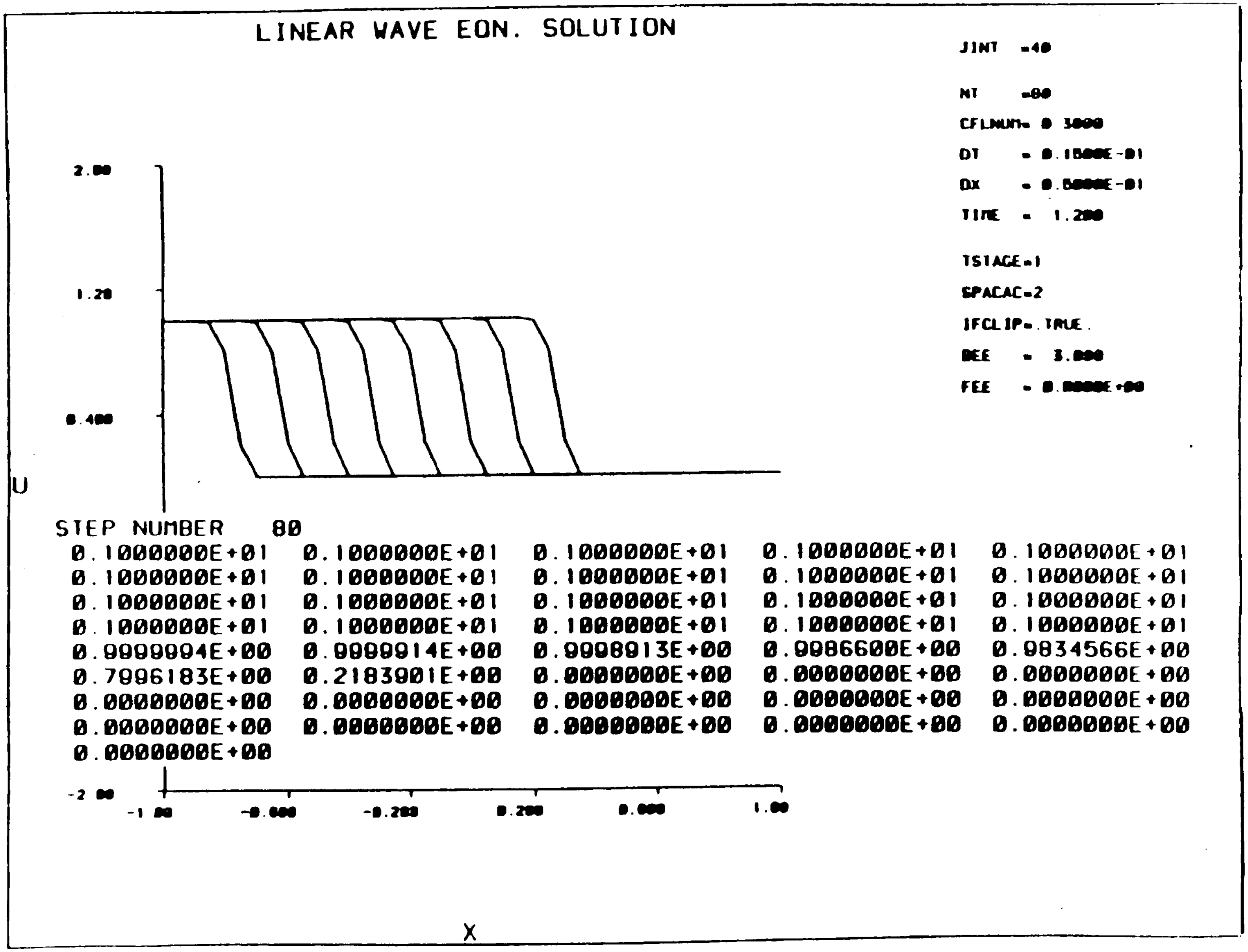
```
JINT   =40

NT     =80
CFLNUM= 0.3000
DT     = 0.1500E-01
DX     = 0.5000E-01
TIME   =  1.200

TSTAGE=1
SPACAC=2
IFCLIP=.TRUE.
BEE    =  4.000
FEE    = 0.3333
```

```
STEP NUMBER   80
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.9999997E+00  0.9999659E+00  0.9973216E+00
 0.8090450E+00  0.1936739E+00 -0.2272492E-08 -0.2956497E-08 -0.1025792E-08
-0.4421143E-11  0.5101777E-17  0.1767341E-17  0.1046846E-19 -0.3984205E-26
-0.2110521E-26 -0.3947276E-28  0.3448769E-35  0.6250161E-36  0.0000000E+00
 0.0000000E+00
```

Figure (7.1a)

Third Order Accurate TVD Solution to $q_t \quad -q_x$

FROMM'S SCHEME

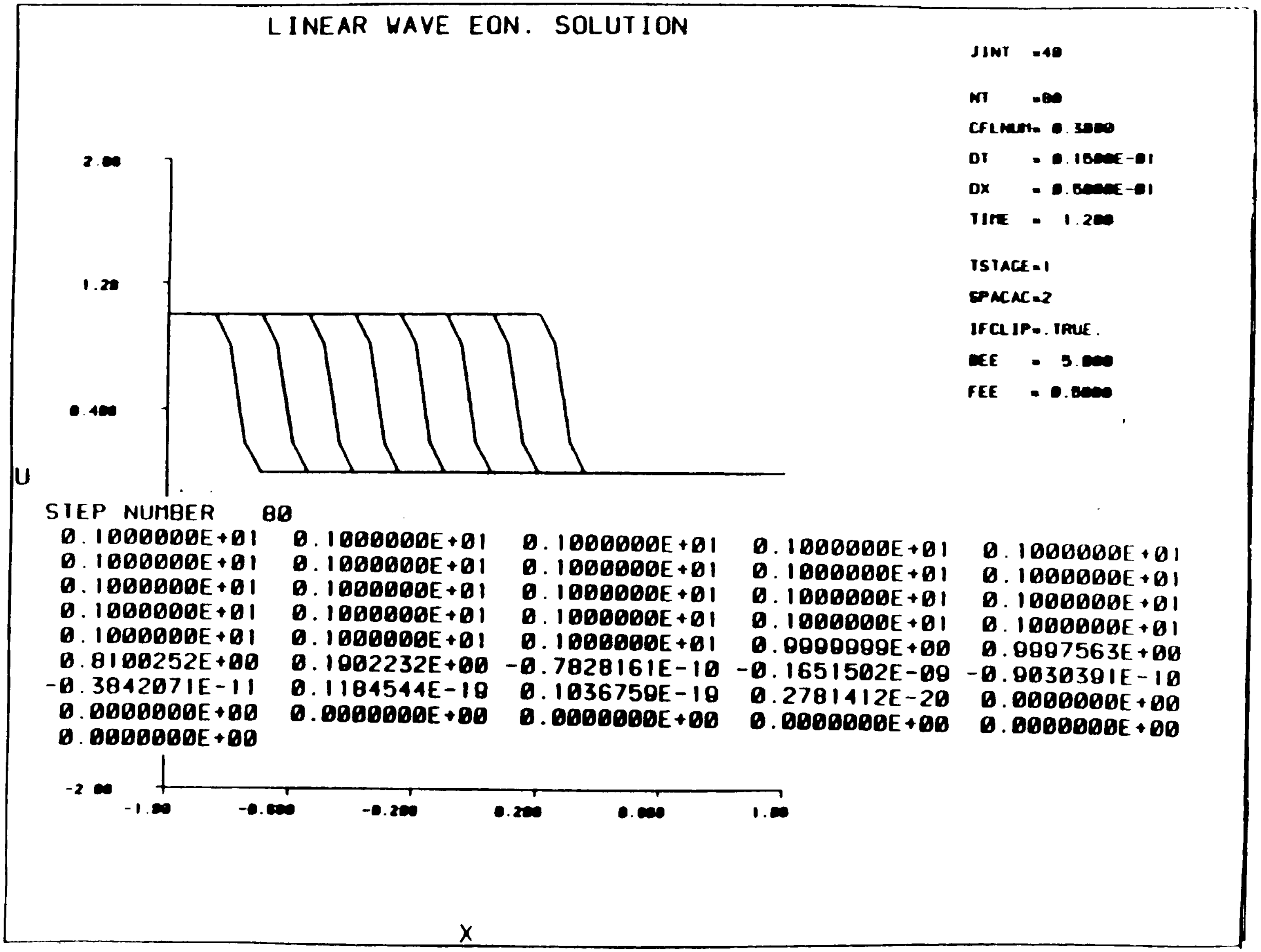


```
JINT   =40

NT     =80
CFLNUM= 0.3000
DT     = 0.1500E-01
DX     = 0.5000E-01
TIME   =  1.200

TSTAGE=1
SPACAC=2
IFCLIP=.TRUE.
BEE    =  3.000
FEE    = 0.0000E+00
```

```
STEP NUMBER    80
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.9999994E+00  0.9999914E+00  0.9908913E+00  0.9086600E+00  0.9834566E+00
 0.7996183E+00  0.2183901E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00
```

Figure (7.1b)

Fromm's TVD Solution to $q_t \quad -q_x$



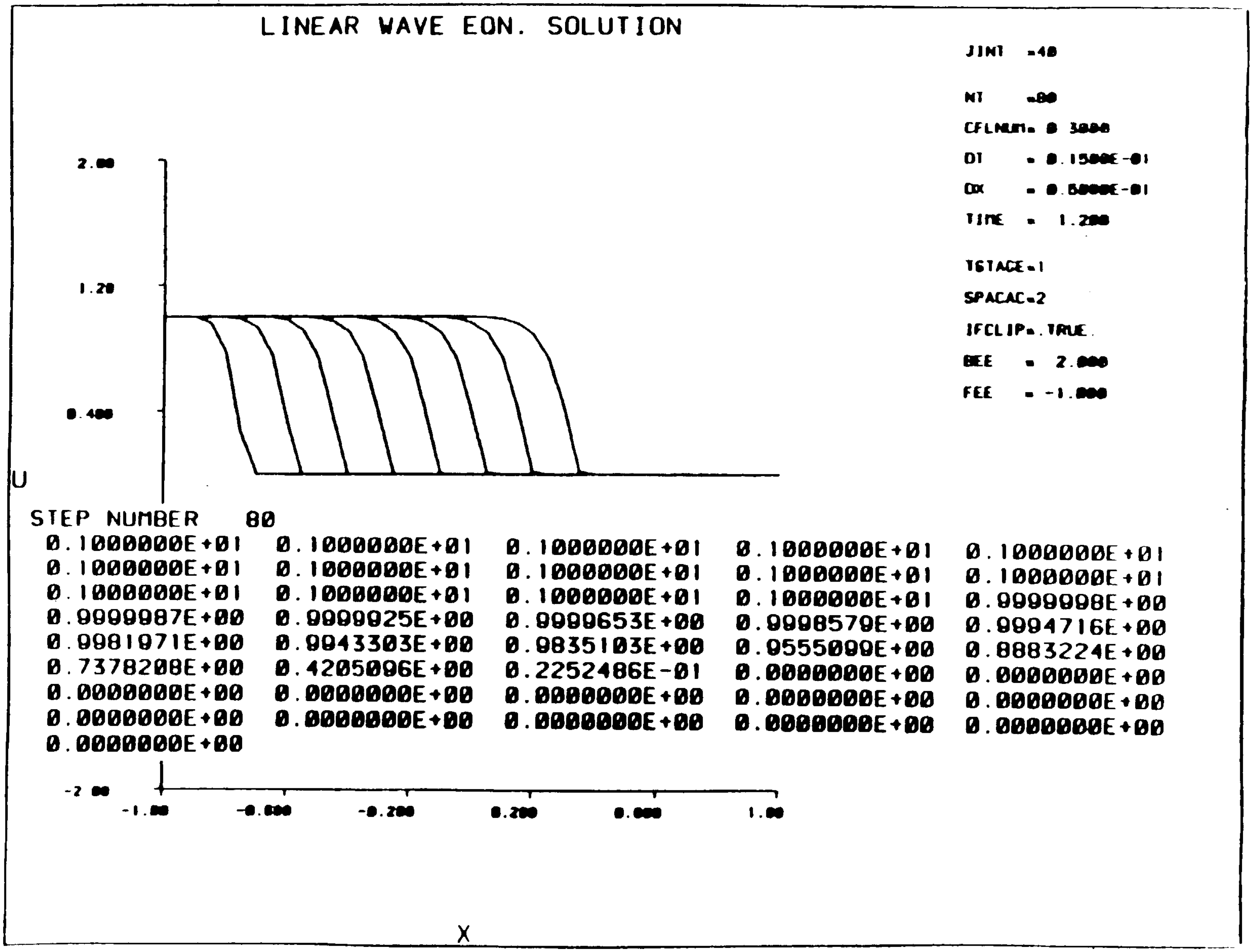
```
JINT   =40

NT     =80
CFLNUM= 0.3000
DT     = 0.1500E-01
DX     = 0.5000E-01
TIME   =  1.200

TSTAGE=1
SPACAC=2
IFCLIP=.TRUE.
BEE    =  5.000
FEE    = 0.5000
```

```
STEP NUMBER   80
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.8100252E+00  0.1902232E+00 -0.7828161E-10 -0.1651502E-09  0.9997563E+00
-0.3842071E-11  0.1184544E-19  0.1036759E-19  0.2781412E-20 -0.9030391E-10
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00                                               0.0000000E+00
```

Figure (7.1c)

High Accuracy Second Order TVD Solution to $q_t = -q_x$

FULLY UPWIND WITH min mod (x,?y)



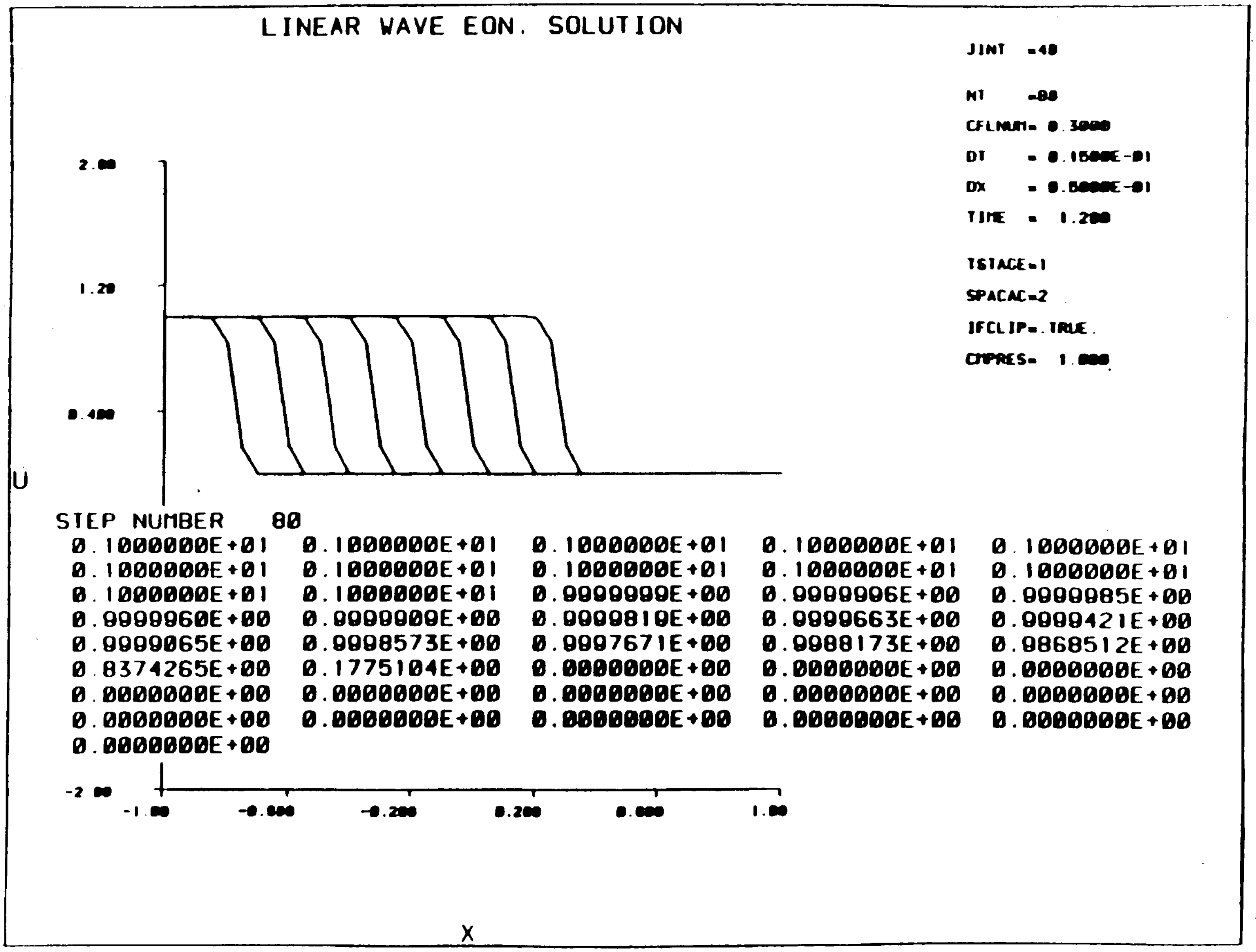
```
JINT   =40

NT     =80
CFLNUM= 0.3000
DT     = 0.1500E-01
DX     = 0.5000E-01
TIME   =  1.200

TSTAGE=1
SPACAC=2
IFCLIP=.TRUE.
BEE    =  2.000
FEE    = -1.000
```

```
STEP NUMBER   80
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.9999998E+00
 0.9999987E+00  0.9999925E+00  0.9999653E+00  0.9998579E+00  0.9994716E+00
 0.9981971E+00  0.9943303E+00  0.9835103E+00  0.9555099E+00  0.8883224E+00
 0.7378208E+00  0.4205096E+00  0.2252486E-01  0.0000000E+00  0.0000000E+00
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00
```

Figure (7.1d)

Fully upwind TVD Solution to $q_t = -q_x$

FULLY UPWIND, WITH SWEBY'S COMPRESSIVE LIMITER

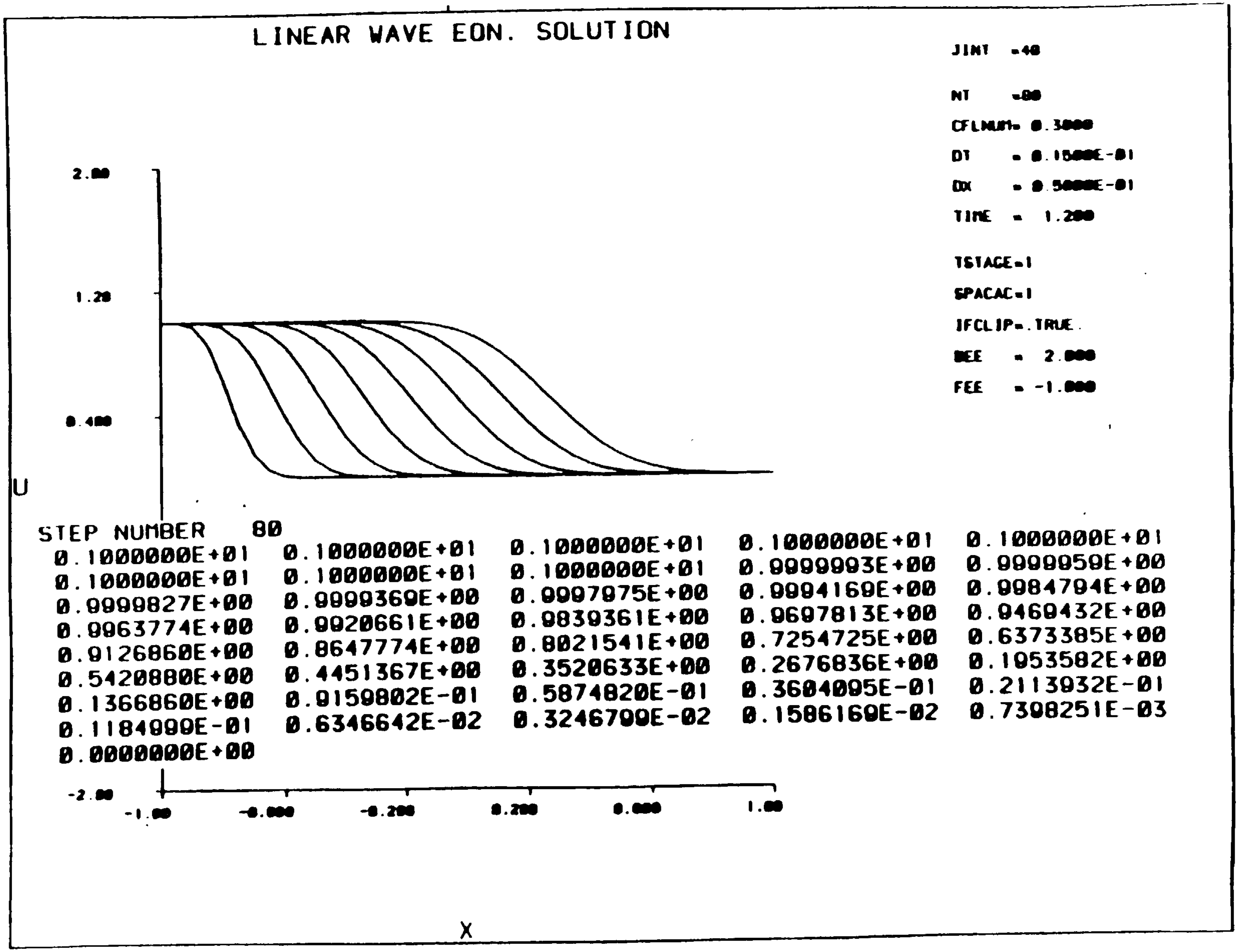


```
JINT   =40

NT     =80
CFLNUM= 0.3000
DT     = 0.1500E-01
DX     = 0.5000E-01
TIME   =  1.200

TSTAGE=1
SPACAC=2
IFCLIP=.TRUE.
CMPRES=  1.000
```

```
STEP NUMBER   80
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.9999999E+00  0.9999996E+00  0.9999985E+00
 0.9999960E+00  0.9999909E+00  0.9999819E+00  0.9999663E+00  0.9999421E+00
 0.9999065E+00  0.9998573E+00  0.9997671E+00  0.9988173E+00  0.9868512E+00
 0.8374265E+00  0.1775104E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00  0.0000000E+00
 0.0000000E+00
```

Figure (7.1e)

Fully upwind, with Sweby's Compressive Limiter, Solution to $q_t \quad -q_x$



```
JINT   =40
NT     =80
CFLNUM= 0.3000
DT     = 0.1500E-01
DX     = 0.5000E-01
TIME   =  1.200

TSTAGE=1
SPACAC=1
IFCLIP=.TRUE.
BEE    =  2.000
FEE    = -1.000
```

```
STEP NUMBER    80
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01  0.1000000E+01
 0.1000000E+01  0.1000000E+01  0.1000000E+01  0.9999993E+00  0.9999959E+00
 0.9999827E+00  0.9999369E+00  0.9997975E+00  0.9994169E+00  0.9984794E+00
 0.9963774E+00  0.9920661E+00  0.9839361E+00  0.9697813E+00  0.9460432E+00
 0.9126860E+00  0.8647774E+00  0.8021541E+00  0.7254725E+00  0.6373385E+00
 0.5420880E+00  0.4451367E+00  0.3520633E+00  0.2676836E+00  0.1953582E+00
 0.1366860E+00  0.9159802E-01  0.5874820E-01  0.3604095E-01  0.2113932E-01
 0.1184999E-01  0.6346642E-02  0.3246799E-02  0.1586169E-02  0.7398251E-03
 0.0000000E+00
```

Figure (7.1f)
First order upwind solution to q_t $-q_x$